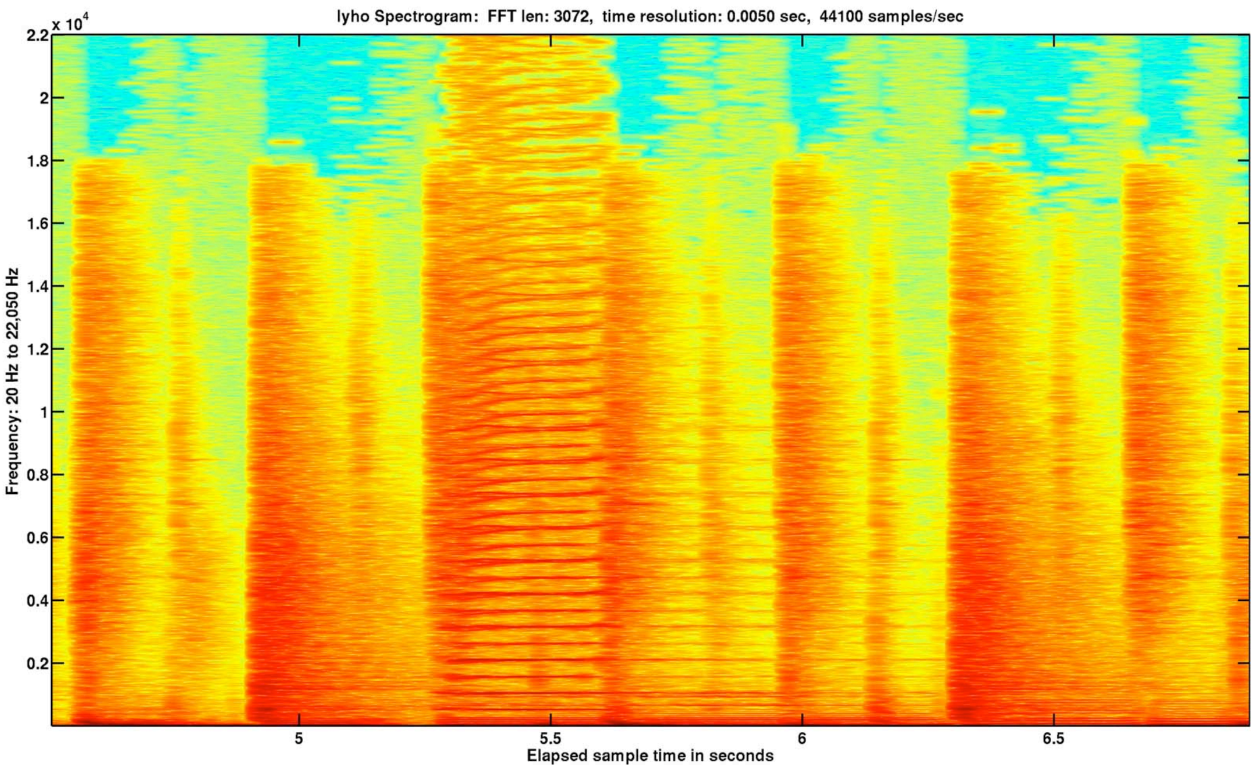
Pulse and Swing:
Quantitative Analysis of Hierarchical Structure in Swing Rhythm



Frontispiece. Trumpet note and drums from You Can Leave Your Hat On *recorded by Tom Jones (1997)*

Ken Lindsay, tlafx, Ashland, OR

Pete Nordquist, Southern Oregon University, Ashland, OR

Pulse and Swing:
Quantitative Analysis of Hierarchical Structure in Swing Rhythm

Ken Lindsay, tlafx, Ashland, OR

Pete Nordquist, Southern Oregon University, Ashland, OR

Abstract

Swing feeling in music is partially attributed to the presence of timing variations in the rhythm as played, contrasted with how it would be written in standard tablature form. Substitution of triplet subdivision for the divide by two format of Mozart-Bach (MB) notation (quarter, eighth, sixteenth notes) is a well known timing variation found in Swing. We present evidence that an important part of Swing is also found in non-uniform timing of the Pulse, or basic beat, of a Swing tune – i.e., time between downbeats of successive musical measures does not conform to a regularly spaced time grid as defined by the metronome.

The departure from a uniform time grid and the presence of different types of time variations played by different instruments both make important contributions to the "lively" quality found in the swing style of many genres: Classic Jazz Swing, Pop, Gospel/Soul, Rock and Roll, Samba etc. Although different rhythms may not be locked together as in MB style, nonetheless they are all highly synchronized. The degree of synchronization corresponds to varying degrees of tight or loose rhythmic style. This also contributes to the "live" feeling in the musical performance. Anyone who wants a direct revelation of the nature of Swing rhythm should take a ride on the the Saint Charles streetcar from the French Quarter out toward Tulane University and Audubon Park.

Prior Research

Last year (2006) we reported our initial findings regarding micro-timing in Swing rhythm [Lindsay & Nordquist, 2006]. Novel features of Swing identified so far include push/pull of note timing around a common subdivision (e.g. ahead and behind a straight quarter note pattern), documentation of non uniform pulse timing, subdivision of the beat by arbitrary numbers (beyond the divide-by-two and triplet formats of MB notation), comparison of tight vs. loose rhythmic style, and interplay between several instruments showing ensemble swing, another source of the spontaneous feeling in live performances, not easily written in standard notation. We have introduced a new visualization form called *diffdot* graphs that show timing details of note events. Diffdot plots easily and clearly document timing subtleties that are fundamental to rhythmic interpretation of music. The triplet notes of Classic Jazz Swing are obvious in the diffdot diagrams on the 1/3 and 1/6 subdivision lines. Triplets (as well as non standard subdivision) are also found in Samba.

Analysis Methods

Our analysis uses straightforward time/frequency methods, which have been described in detail elsewhere [Lindsay & Nordquist 2007]. Figures 1 & 2 show the process of transforming musical audio data through a fairly standard time/frequency *spectrogram* form, and thence into time and note event graphs. Audio power levels are indicated by color in the spectrograms, and this information is converted to time series format showing the time changing power level (loudness) in the music. The two types of event graphs are 1) multi-frequency time series plots of audio power in each frequency band, and 2) *diffdot* plots that show note events as both chronology (like the time series plots), and note time durations or *inter onset intervals* (IOI) as they are sometimes called. It is the diffdot plots which give the most direct insight into the subtle features in the music that give rise to Swing feeling.

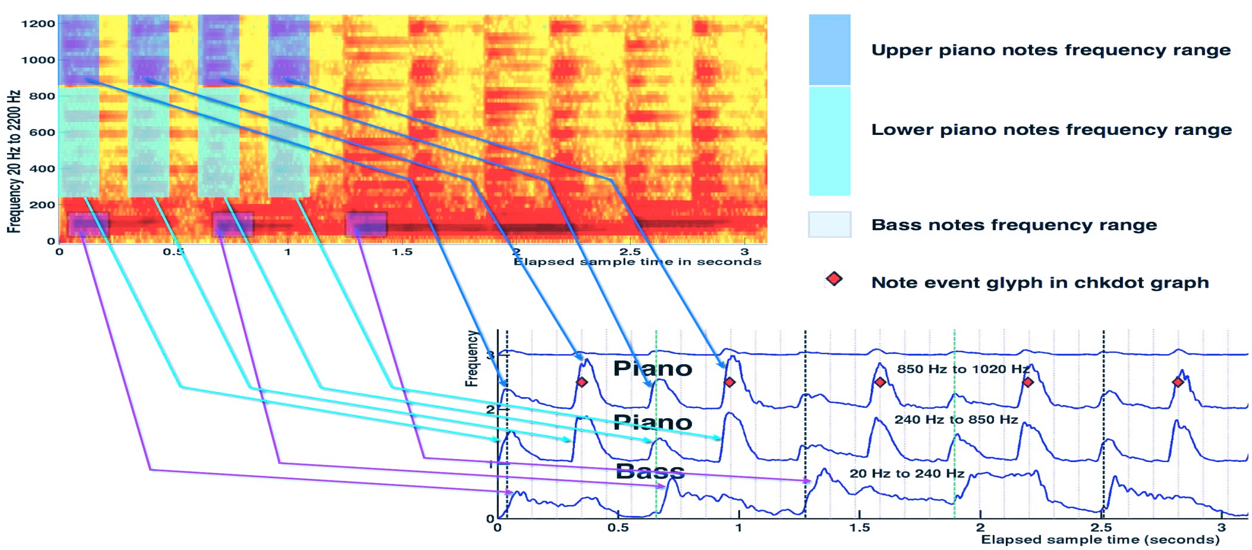


Figure 1. Transforming musical note information from spectrogram to time series form.

Diffdot plots show the actual time location of each musical note along the X axis. Vertically we plot the time difference between successive notes in each rhythmic group (e.g. pulse and secondary events). By horizontal markings on 1/2, 1/3, 1/4, 1/6 and 1/8 subdivisions of the pulse, we indicate where standard musical notes would be graphed. If the pulse is a whole note, then the 1/2 subdivision beat line is a half note, 1/4 is a quarter note and 1/8 is an eighth note. The 1/3 and 1/6 lines represent *triplet* subdivision which is a common feature in many but not all Swing styles. We often mark the pulse using backbeat note events, played by the hi-hat cymbal for example, which is usually considered a half note. In that case the 1/2, 1/3 etc. beat lines would be note time durations of half the subdivision, e.g. the 1/2 subdivision beat line is a quarter note. The relative timing patterns between notes remain the same when counting the pulse either as a whole note or half note. In more complex rhythms such as *Iko Iko* (the Bo Diddley beat, related to standard Cuban Clave), or Brazilian rhythms, these subdivisions can become a bit more complicated.

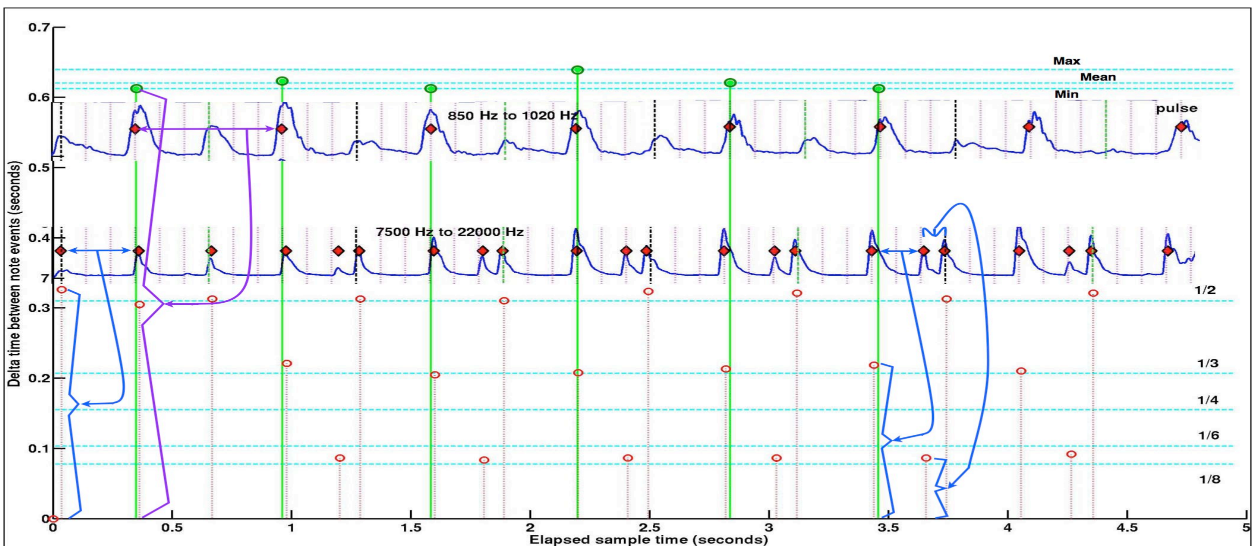


Figure 2. Transforming musical information from time series form into diffdot notation.

Musical Examples

We first analyze the Swing found in the simplest rhythm of all: the *backbeat*. Familiar to most people as the evenly timed *boom-chuck!* beat that is the foundation of many forms of pop and dance music, this rhythm is straight forwardly counted as *1*-**2**, *1*-**2** or ***1***, *2*, ***3***, *4* or *1*, ***2***, *3*, ***4*** and similarly (bold typeface indicates a musical accent), alternating the downbeat and backbeat. In the real world of music performance, there is a lot more going on here than merely repetitious counting to 2 or 4. This is perhaps the most important fundamental conclusion supported by our current evidence: complexity, not clockwork, helps create Swing.

Let us begin by examining the musical pulse (or fundamental beat) in ***Duke's Place***, a 1962 recording by Duke Ellington and Louis Armstrong. Figure 3 shows a 5 frequency-band time series plot of a short section of the tune. In the fourth row from the bottom, the pulse of the music is plotted. In this case, the hi-hat cymbal plays a constant ride on the backbeat which is the most consistent beat in the music. Hence we use this instrument to mark the beat and meter since it is closest to standard MB counting. Four bar phrases are marked by vertical black lines, and subdivided by green lines. Each single bar, or measure, is further subdivided by 6, which lets us show both the downbeat and backbeat, as well as any triplet subdivision notes which are so common in Swing rhythm. The uniformity of the vertical subdivision lines parallels the uniformity of similar markings in sheet music. Unlike sheet music however, the note events themselves are not constrained to lying on the standard subdivision of whole note, half note, quarter note etc. We plot the notes with their associated high resolution time locations as they happen in the audio stream of the recording.

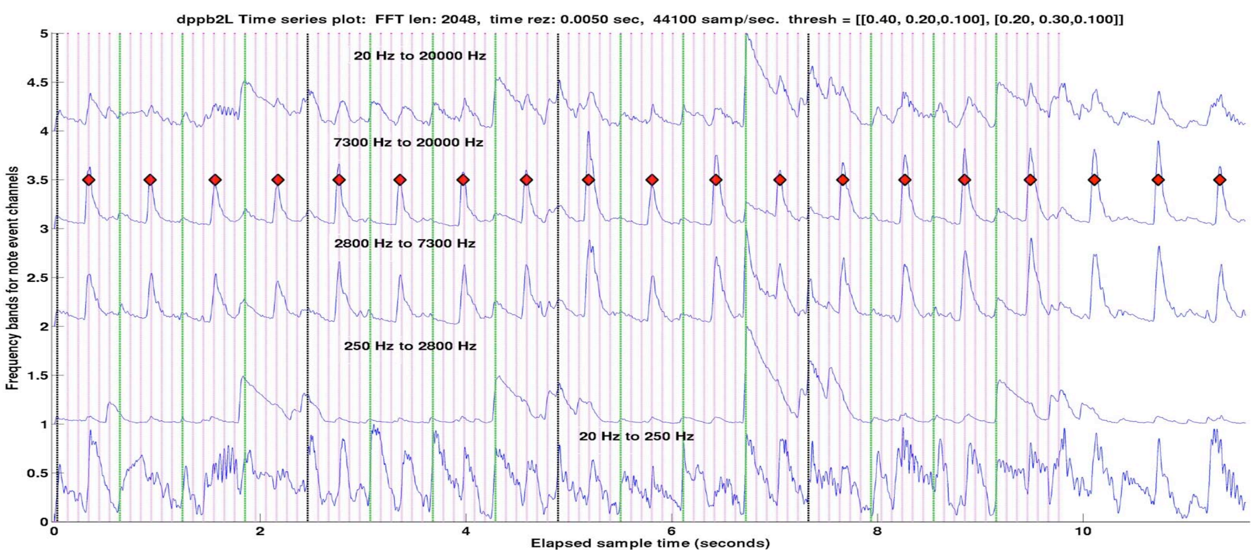


Figure 3. Pulse beat in Duke's Place *recorded by Duke Ellington and Louis Armstrong (1962).*

To a casual eye, the note events marked in figure 3 seem quite uniformly spaced in time, all landing on the backbeat lines. Indeed the deviation of note event times from where they officially "should be" (from the MB view) is fairly small, on order of a few tens of milliseconds. By looking at the time deltas between note events, we discover a microcosm of order and pattern within these short time variations that correspond to musical note durations of 1/32 and shorter.

Figure 4 shows the same information in diffdot form. The variations of pulse note timing are obvious. The pulse markers in the diffdot plot wander up and down fairly smoothly, but without any particularly obvious macro-pattern. The abrupt short/long pair after the 8 second mark corresponds to a note duration between a 16th and a 32nd note. Thus the maximum range for pulse timing variation during this section is less than 600 milliseconds (average time for the half note backbeat) divided by 8 (mean pulse time multiplied by 1/8th note duration), or about 75 milliseconds. Most of the pulse time deltas are less than half of this figure, 35 milliseconds or less. This represents a fairly loose rhythmic style.

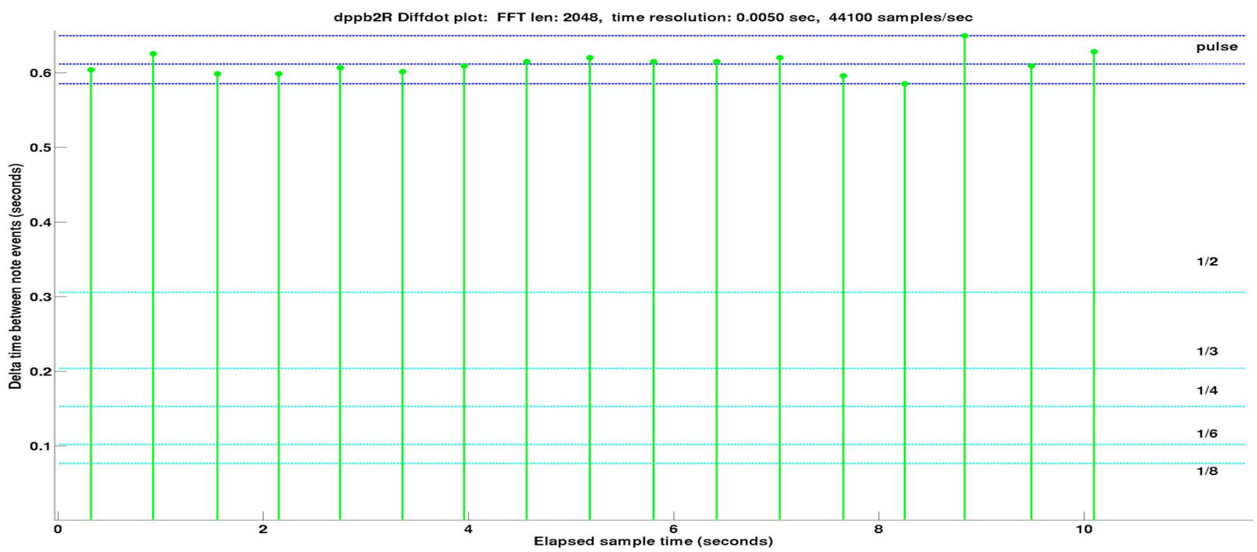


Figure 4. Diffdot plot of pulse beat in Duke's Place *(1962).*

The middle row of figure 3 shows some slight bumps corresponding to the walking bass line playing notes twice as fast as the hi-hat (i.e. quarter notes). This is shown as diffdot form in figure 5, although we miss a few note events because the bass is not very loud and is partially masked by piano and drums. Nonetheless, the timing patterns of the secondary events show some interesting features that are not seen in the pulse. In the right half of the diffdot plot, most of the pairs of bass notes show a short/long pattern with the shorter note anchored on the green line, locked into the main beat of the hi-hat. Thus the combination of bass and hi-hat notes give a non-uniform feature in the timing that could be notated precisely in terms of 32nd or 64th notes but more likely would be written loosely as whole, half or quarter notes. The sheet music would generally include "with a swing feel" or similar instruction to indicate a departure from strictly evenly timed MB style.

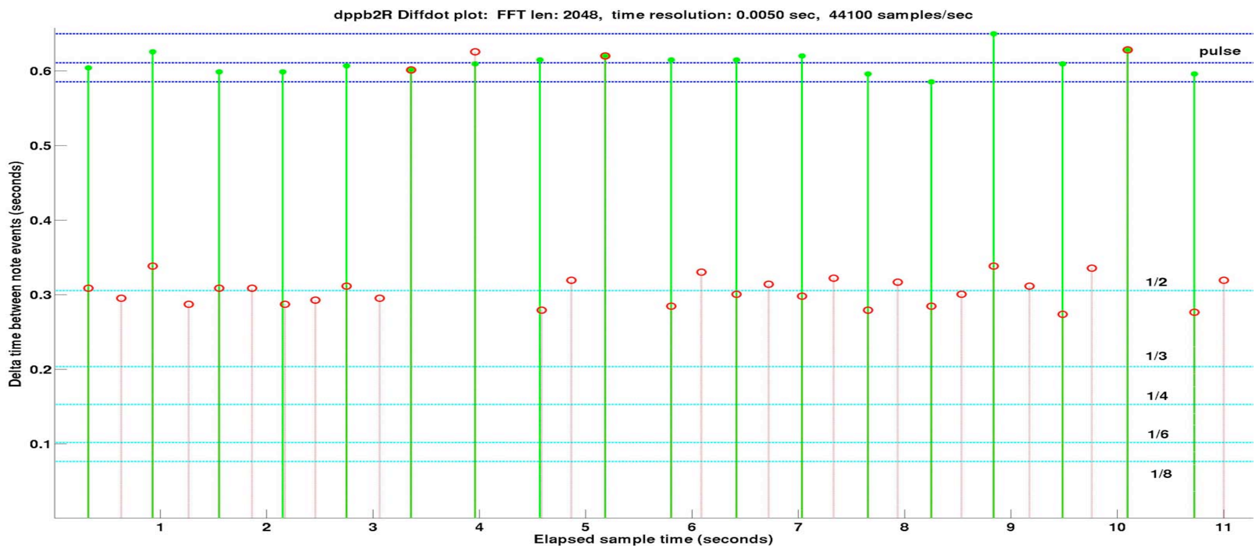


Figure 5. Pulse (hi-hat) and backbeat (walking bass line) note events for Duke's Place (1962).

Finally in figures 6 and 7, we show the time/frequency spectrogram of the data. Hi-hat cymbal is visible in the top half of figure 6. Duke's piano is found in the striated blocks extending to 4 KHz around the 2, 5, 7 and 10 second marks. The red swath at the bottom is the walking bass line and drums. Figure 7 shows a close-up of the low frequencies.

Duke's piano can also be seen in the second row of figure 3. Close inspection of the time series plot (figure 3) shows that there are several small events that precede the major shapes of the graph in the second row which are the piano's sound. These events are typically on the triplet subdivision line pickup to the main beat (pink lines immediately preceding black lines). This leading triplet note is a common motif in American Jazz and Swing music. There are also some piano events that happen on the second triplet subdivision line following some of the main beats, e.g. around the 5 and 7.5 second locations.

Duke Ellington's rhythmic style includes timing variations that are close to 10% of the pulse, on the order of 40 to 50 milliseconds. Because the recording is live, and was done before the days of micro editing of audio track timing in the post-production engineering process, we conclude that such micro timing is an inherent facet of performing classic Jazz Swing. The presence of triplet subdivision features are also typical of this musical style.

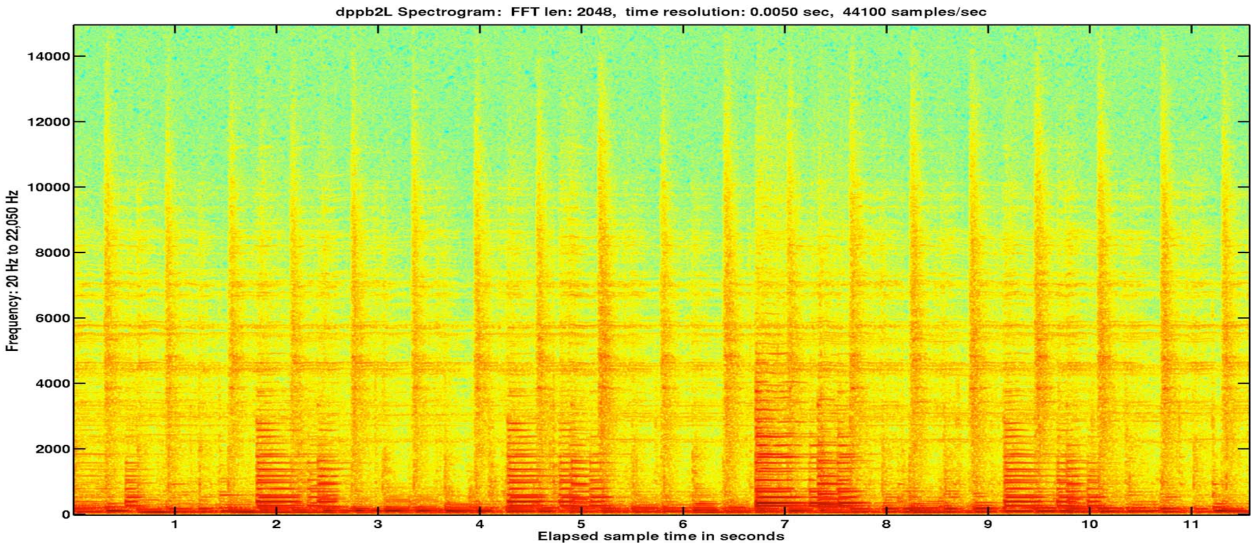


Figure 6. Spectrogram of a 12 second sample of Duke's Place (1962).

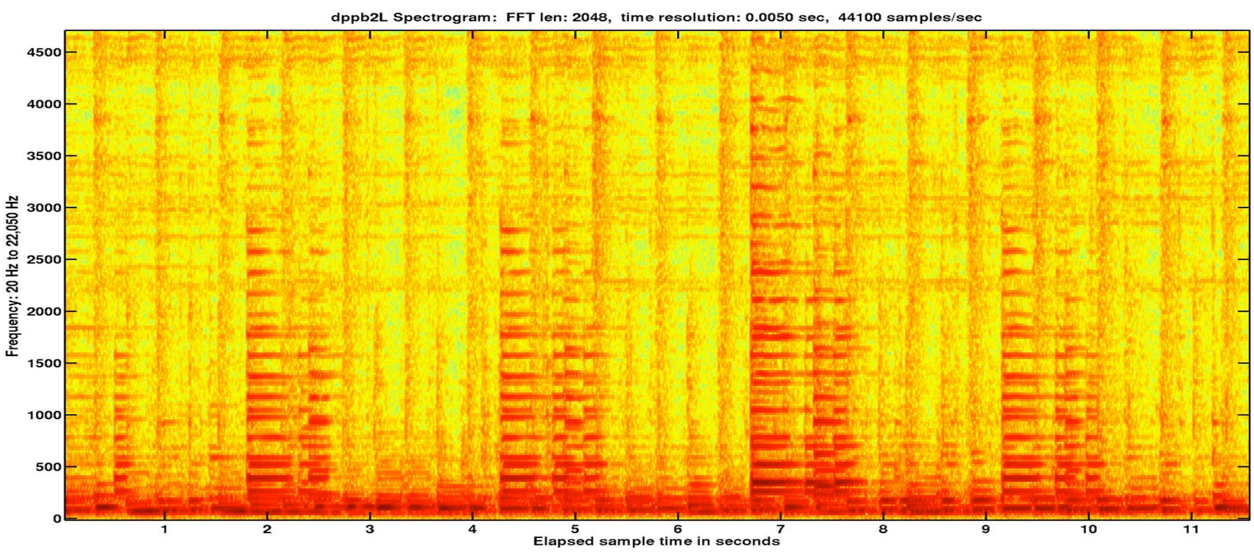


Figure 7. Close-up of low frequencies in spectrogram of Duke's Place (1962).

We next examine several more modern examples of backbeat which come from the broader field of Pop music: *I Heard it Through the Grapevine* (Motown), *Hit the Road, Jack* (R & B), and *You Can Leave Your Hat On* (Dance Hall Burlesque). None of these tunes has any obvious triplet subdivision, but they all meet our fundamental measure which indicates the presence of Swing style: the rhythm makes us tap our foot, bounce our body, or get up and dance. Additionally, they meet our second criterion for Swing: a short loop can play endlessly and does not become tedious after a few repetitions.

I Heard it Through the Grapevine is a fairly typically Motown tune with a strong and straightforward fundamental beat, having more complex rhythmic layers weaving in and out on top of the basic beat. We only examine the fundamental beat structure here. The bass drum plays a *ba*-***bump!***-*ba*-***bump!***-*ba*-***bump!*** heartbeat, and the hi-hat cymbal plays the straight backbeat counterpoint. The diffdot form shown in figure 8 reveals that these simple beats are readily described using whole notes and quarter notes. However, listening to the recording, there is an insistent and passionate quality to the rhythm which is common in the emotionally powerful Motown style. Straight sheet music and drum machines would probably have difficulty capturing this quality.

Several features are readily apparent in figure 8 which suggest a logical source for the emotional effect. Although the pulse has a wandering quality reminiscent of *Duke's Place,* the note events lying on the 1/4 subdivision line (eighth note pickup in the *ba*-***bump!***), occur in pairs which have a slight long/short pattern. This and the inverse short/long pattern occur in several other tunes we will analyze shortly. There is the obvious viewport issue: does such a pattern have a long/short or short/long emphasis? This can be determined by examining other information such as the relative accent (loudness) of the two related beats, or which beat happens on a pulse beat (neither in this case).

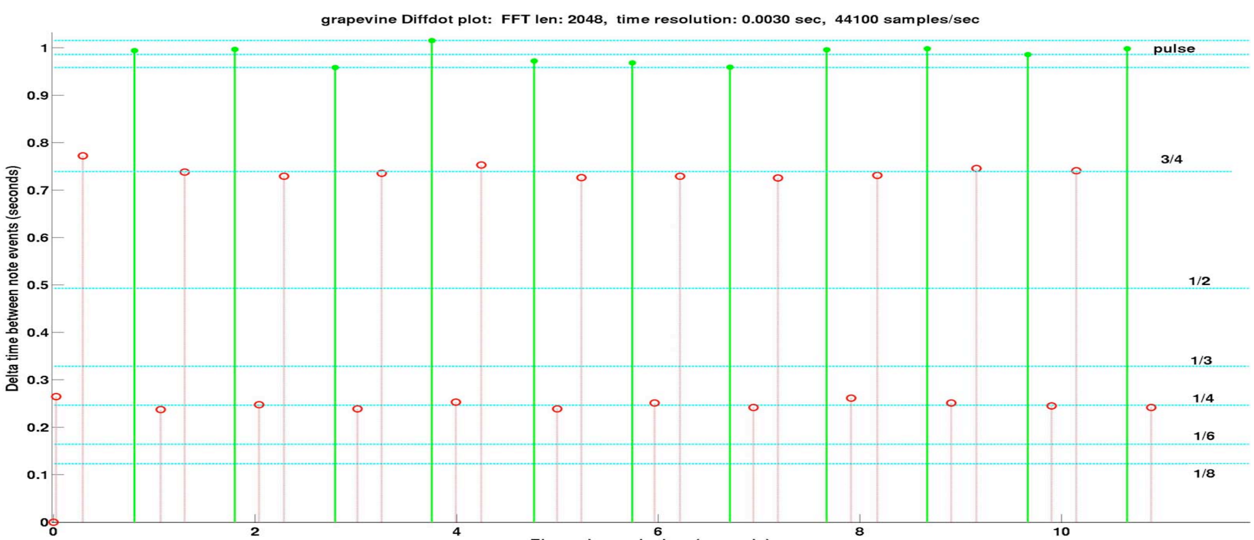


Figure 8. Backbeat and heartbeat note events in I Heard it through the Grapevine (2002).

The note events on the 3/4 subdivision beat line in figure 8 represent the long pause between one *ba*-***bump!*** and the next one, corresponding to three standard MB quarter notes.

In this example, the rhythmic figure would be justifiably written using straight quarter note subdivision in MB notation. Playing the sheet music through a drum machine or computer sequencer would give a sterile and fairly uninteresting recording by comparison to this one played live by the Funk Brothers. An interesting subtlety in this version can be seen by connecting the dots on the 3/4 subdivision line: broad curves which are concave upward emerge by joining several beats together. This is a slower event structure than the long/short pattern observed on the 1/4 subdivision line. While we have not examined the entire recording, it seems likely that this non random "breathing" feature would be found commonly, although perhaps not present throughout the entire song. This is a topic for further research.

Time variation for pulse note events is less than 45 milliseconds. The time differences between note events on the 1/4 subdivision are less than 20 milliseconds. Note events on the 3/4 subdivision are generally within 25 milliseconds, although the difference between the first two note events are approximately 40 milliseconds. Counting the pulse as a half note, a 32nd note would be around 60 milliseconds, so all these timing variations are in the 32nd to 64th note range, or quicker.

Figure 9 shows the time series plot of the note events from figure 8. Slight variations can be read directly by whether a red diamond note event is exactly on the vertical subdivision line. The metrical structure is taken from the hi-hat backbeat. A slight tempo increase is obvious as the note events on the right side of the plot creep ahead of the subdivision lines. The time locations of the subdivision lines are set by the first four backbeat note events shown on the fourth row up. Clearly this calls for a code enhancement which adjusts the subdivision lines adaptively to the slight changes in the beat, rather than making a rigid unchanging framework as is done here, and as is inherent to the meter and subdivision of MB notation.

The issue of tempo change can raise some subtle issues when analyzing swing styles. Generally, tempo is considered to be a long term statistical average of many pulse beats, but during an accelerando for example, the changing time delta between every two beats would be important. Other non random beat time fluctuations can be considered part of the rhythmic style, as in the breathing feature mentioned above.

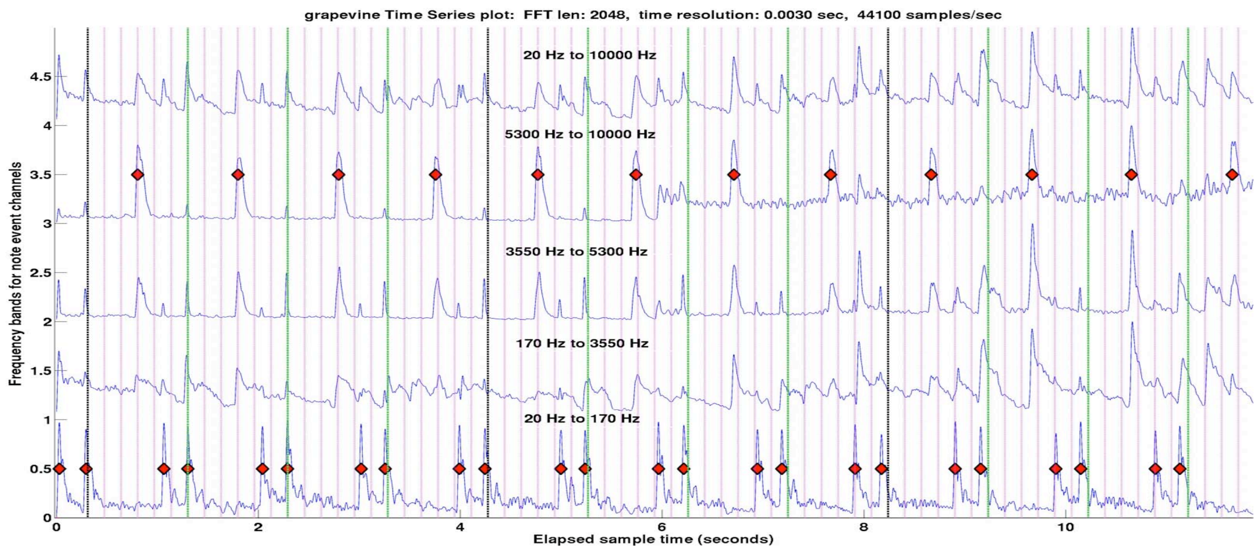


Figure 9. Time series showing heartbeat (bottom) and backbeat (4th row up) for Grapevine *(2002).*

Figures 10 and 11 show the spectrogram for *Grapevine.* Figure 10 shows the low frequencies from the bass drum *ba-**bump!*** note events. This figure is a good example of how some frequencies in a note event seem to happen at slightly different times. While this is mentioned occasionally in the literature, we have not seen a good mathematical analysis of what it means. It may be an artifact of Fourier methods, or may exist in the audio itself. In figure 11, the hi-hat creates the twelve tall fat red columns, while the bass drum double *ba-**bump!*** beats are visible as sharp thin vertical red lines in the 3 to 8 KHz range. Starting at the 6 second mark are high frequencies caused by the jingling of a shaker type instrument.

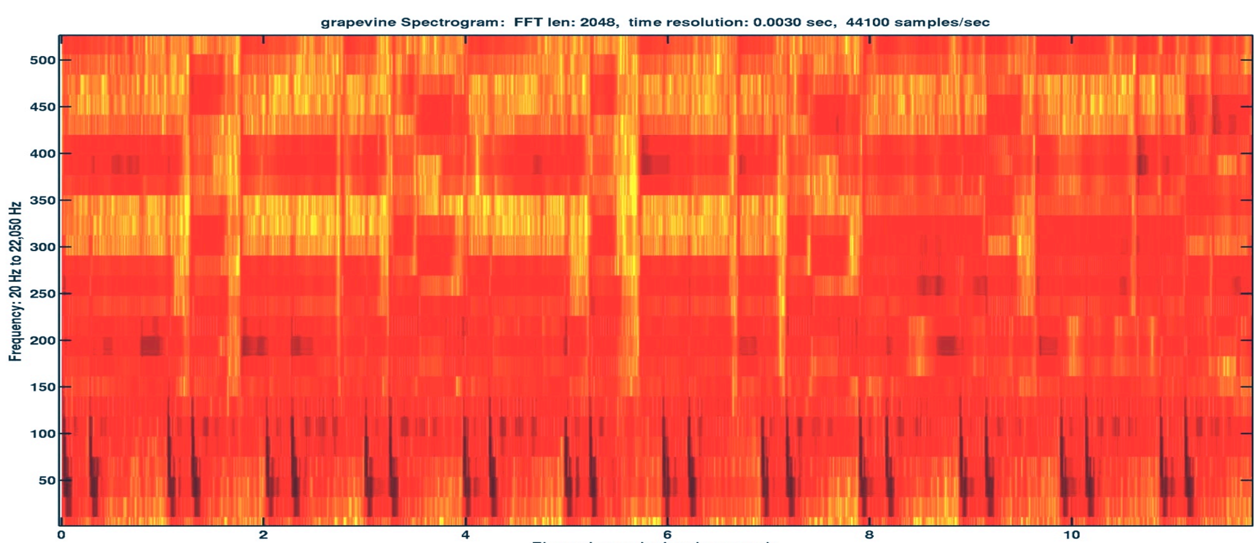


Figure 10. Spectrogram showing close-up of heartbeat for Grapevine *(2002).*

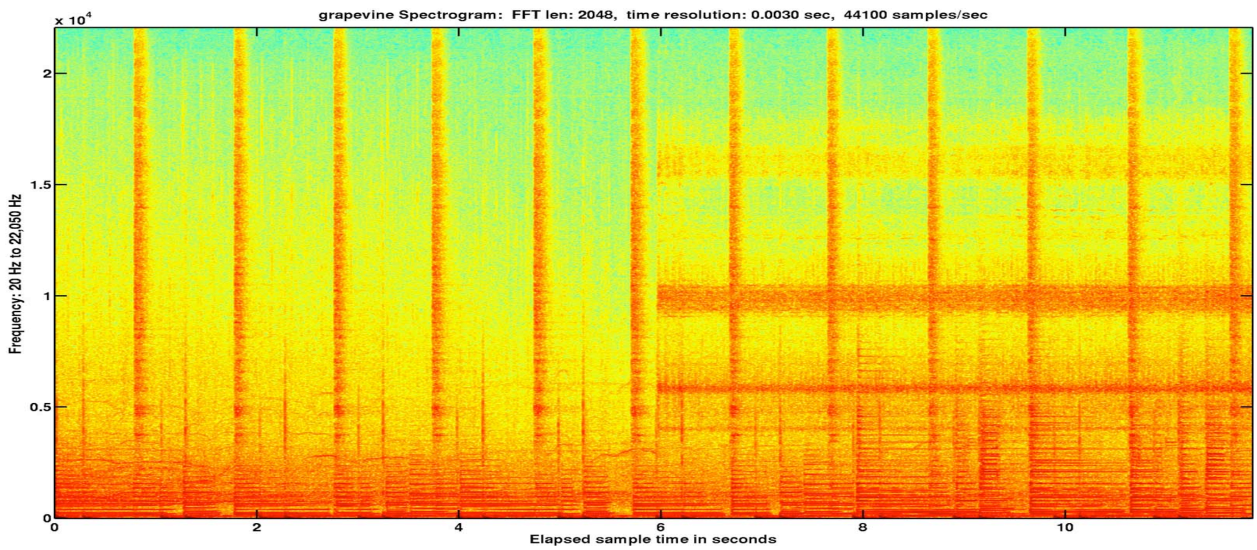


Figure 11. Spectrogram showing backbeat and heartbeat for Grapevine *(2002).*

Hit the Road, Jack! is a Ray Charles tune written by Percy Mayfield which crosses several musical genres. Not being musicological experts, we have called it R & B, but others may classify it differently. Figure 12 shows the pulse and backbeat foundation of this tune. Figure 12a shows a larger view which includes the rhythmic break at time 4.5 seconds.

Notice that the backbeat note event preceding the rhythmic break is substantially higher than the mean pulse line (average of all pulse events). This indicates a slight *ritard* as part of the break: i.e. the beat timing slows briefly at this point in time to emphasize the rhythmic break. While not intrinsically a feature of Swing, *ritard* is a common feature in many types of music, and is clearly documented by diffdot notation. Similarly, the green pulse event preceding the rhythmic break lands on the vertical axis near the 1.5 second delta time point. This is about 10 to 15% longer than if the beat continued uniformly, without the *ritard*. By including this long pulse beat in the calculation of the mean pulse time, the ostensible beat timing is increased slightly. An artifact of this is that the beats on the 1/2 subdivision line are slightly lower than the 1/2 line. The 1/2 line is slightly *too high* because the pulse line is also slightly too high from the long pulse beat note. The slight flaws in annotation details do not change the overall patterns of the Swing features we are inspecting.

In contrast to Duke Ellington's style, Ray Charles' pulse shows a consistent short/long pattern throughout. The last few seconds have a more uniform pulse, so the pattern documented here may change as the song transitions from the intro to the main part. The 1/2 beats show an alternation between short/long and long/short patterns, discernible by noting whether the short or tall marker is on the green line. There is a fair amount of regularity, but the pattern is far from mechanically repetitious. It is likely that these features, both regular and quasi-random, enhance the swing feeling of this song's rhythm.

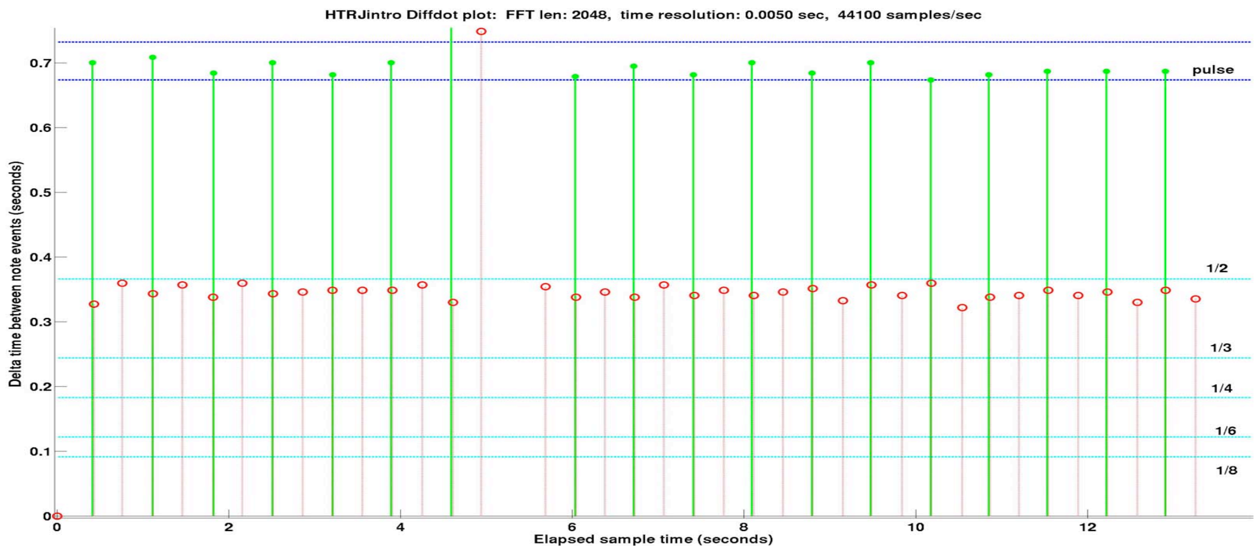


Figure 12. Diffdot showing pulse (top) and backbeat (middle) for Hit the Road, Jack! *(1961)*

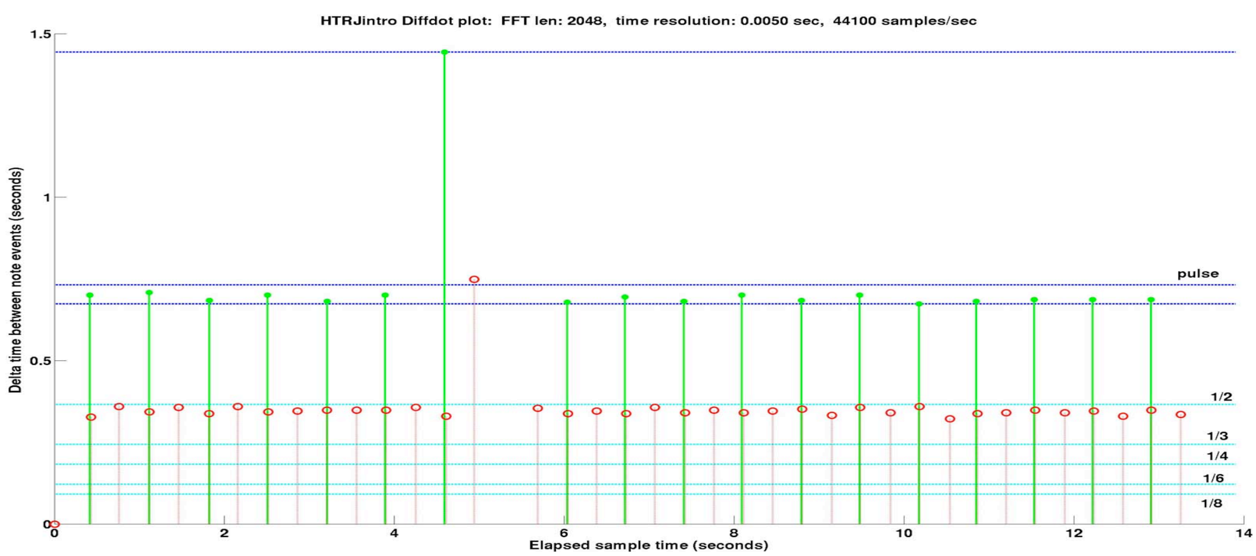


Figure 12a. Diffdot plot showing pulse, backbeat and rhythmic break for Hit the Road, Jack! *(1961)*

Each short/long pair of note events in the pulse has a time difference between about 15 and 25 milliseconds. The *total* variation is around 40 milliseconds since the pairs do not have completely uniform timing between them. This represents a fairly tight rhythmic style by itself, in addition to the well defined Swing features. The total range of 1/2 beats is similarly around 40 milliseconds, and the short/long and long/short pairing are also closer in time than the overall range of time variation for this group of note events. These time deltas represent note durations of less than a 32nd note for the Swing features. During the short sections of negligible time delta variation, the beats are as tight as a mechanical metronome. It is likely that Ray Charles used the alternation between a straight feel and a short/long Swing motif as an important part of changing the mood within the song from moment to moment.

Figures 13 and 14 show the time series events and spectrogram respectively. Unlike previous examples, the beat is very consistent over the course of time for this sample. There are no signs of tempo variation such as the creeping of note events relative to the subdivision lines in the time series graphs.

In this one tune, Ray Charles has included a tight rhythmic style, two types of Swing detail (short/long and long/short), alternation between Straight Time style and Swing, as well as a slight *ritard* and rhythmic break. We have not found any triplets, which some people consider the hallmark of Swing, but we assert that this tune does indeed Swing, based on our two Swing criteria: 1) makes people tap their foot, dance, or move their bodies, often fairly unconsciously; and 2) a short loop of the recording can repeat indefinitely and does not become rhythmically tedious. .

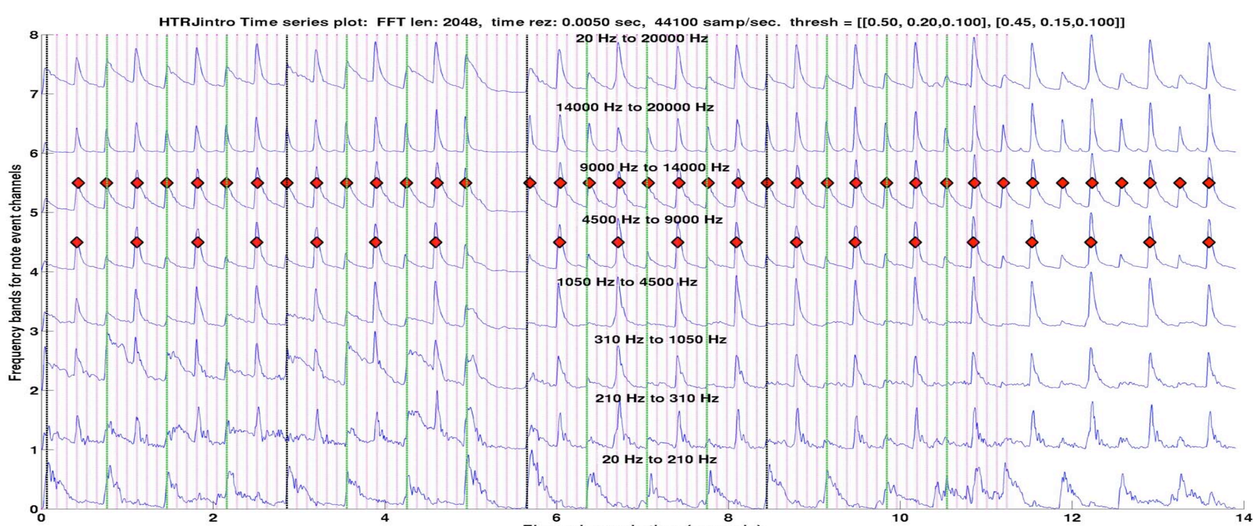


Figure 13. Time series showing pulse, backbeat and rhythmic break for Hit the Road, Jack! *(1961)*

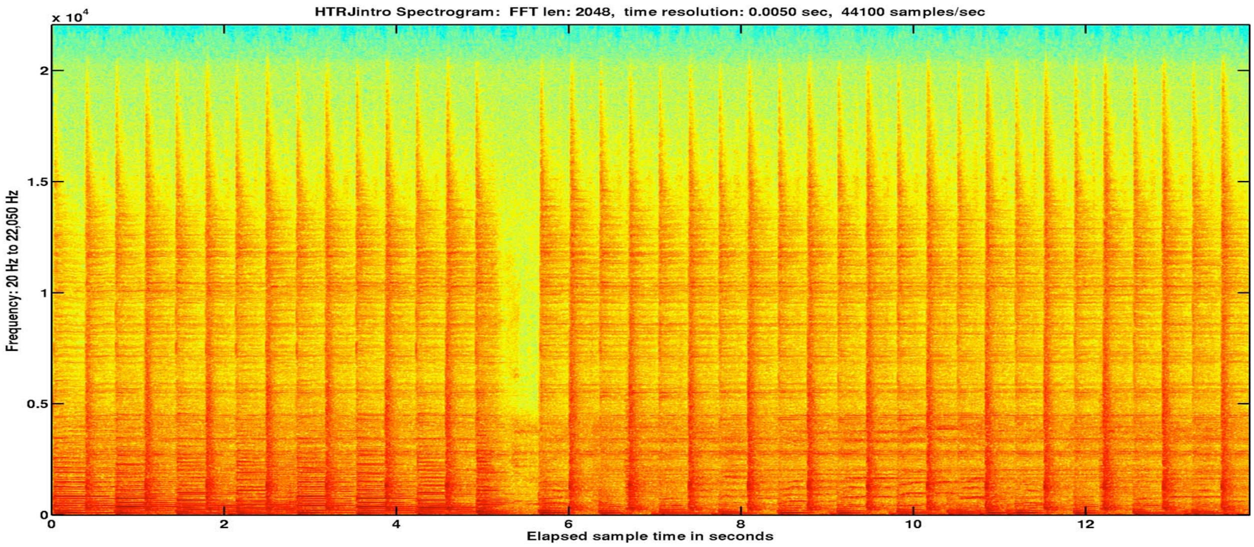


Figure 14. Spectrogram showing pulse, backbeat and rhythmic break for Hit the Road, Jack! *(1961)*

Our final backbeat song in this paper is ***You Can Leave Your Hat On*** written by Randy Newman (1975) and recorded by Tom Jones for the film *The Full Monty* (1997). A bump and grind tune of juicy proportions, this recording has a different type of bouncy feeling than the previous examples. Figure 15 shows the pulse (4th row), and secondary beats (1st row) which all line up on some variation of quarter notes.

The first thing to notice in the diffdot plot, figure 16, is that the pulse time variations are much less than in the previous examples, i.e. a *very* tight rhythmic feeling. Closer inspection of the data has shown the time range for pulse time variations to be somewhere between 15 and 20 milliseconds. Both *Duke's Place* and *Grapevine* have pulse time ranges around 40 or 50 milliseconds. Pulse times for *Hit the Road, Jack!* showed a micro structure in the 25 millisecond range, and an overall variation in the 40 millisecond range.

The secondary beats in *Leave Your Hat On* all fall on the 1/2 and 1/4 beat lines. These are the note events in the bottom row of figure 15. The groups of three in figure 15 are transformed into the groups of two in figure 16, which shows the time *differences* between notes. The time *between* groups of three show up in figure 16 on the 1/2 line. As usual, the pulse is shown along the top. The extreme precision of this recording has led us to wonder if any part of it is from a drum machine. This precision has also revealed some important technical issues with respect to our use of time/frequency granularity -- the time measurements in this tune were sometimes garbled by data precision artifacts, necessitating a finer grained analysis than we use for many tunes. In general we have used 5 millisecond precision for these examples, but for *Leave Your Hat On* we also used 2 milliseconds. It would have been useful to have finer time resolution, but we ran up against software memory limitations.

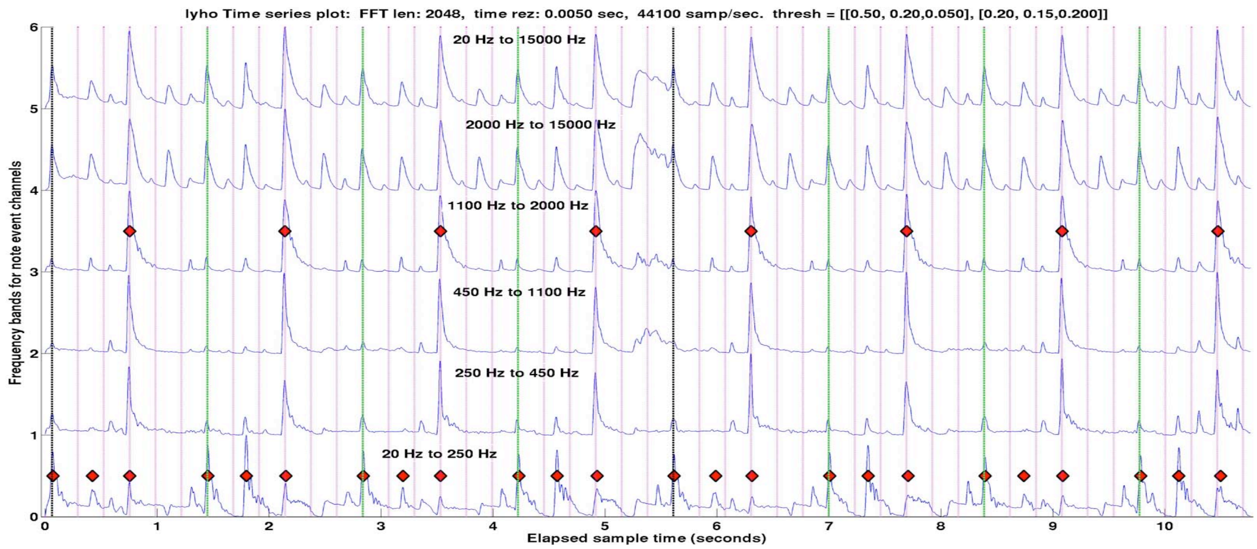


Figure 15. Time series plot of pulse (row 4) and secondary beats (row 1) for You Can Leave Your Hat On.

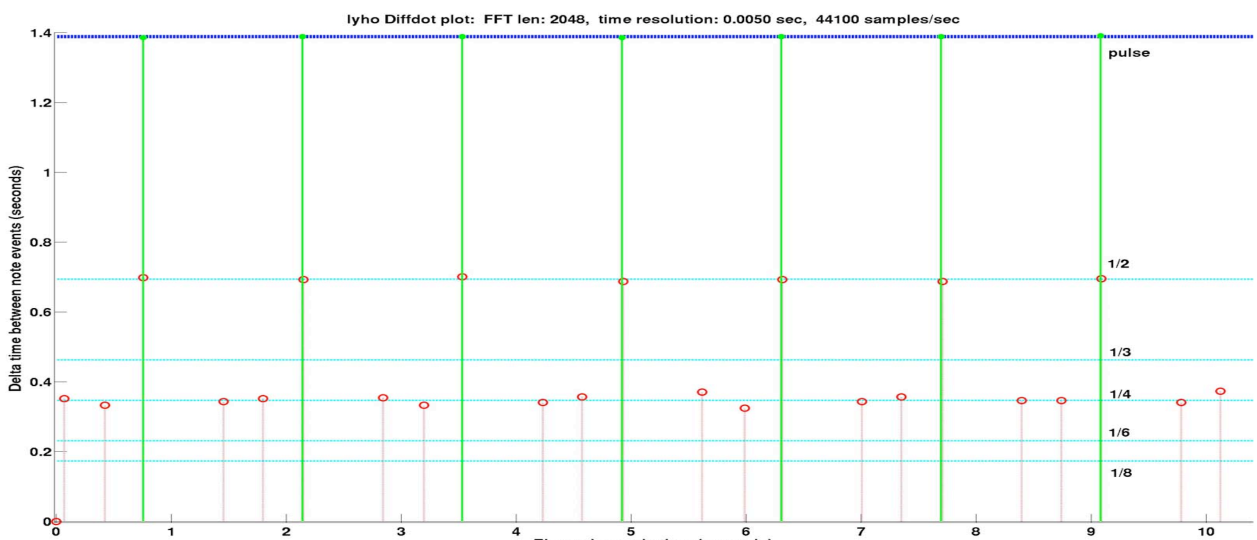


Figure 16. Diffdot plot for You Can Leave Your Hat On *(Tom Jones, 1997)*.

The 1/2 beat events are almost as uniformly timed as the pulse. The 1/4 beat events show a variation of the previously observed long/short patterns. Long/short alternates with short/long in a very consistent manner throughout this sample. There are some slight non uniform variations in the exact parameters of the short/long timings, which indicates rhythm played by a human rather than a computer. The tight pulse is the primary suspect of being played by computer, whereas the other note events have these human quasi-random variations. We have recently received some information from Steve Schown (software engineer and MIDI musician) regarding the possible precision and patterns of note timing artifacts in rendered MIDI files (a data granularity issue), so the question of non-human rhythms in this tune is still open. Personally, we think it is a really good human drummer, primarily from observing by ear the subtle variations of acoustical dynamics in the recording.

Like the previous examples, the short/long and long/short features in figure 16 indicate the presence of Swing feeling in the apparently complete absence of triplet subdivision. The more consistent repetition style here contributes greatly to the tight rhythmic feeling while maintaining the Swing feeling.

In figures 17 and 18 we use slightly different note detection parameters, with the same audio sample as our source data set. The new parameters detect some notes in figure 17 which are between note events found in the bottom row of figure 15. We plot the time deltas in the diffdot figure 18. At the 5 second mark we miss one note event which makes this section appear different from the first three groups. We also miss a note event at 7.5 seconds which changes the pattern there as well. Nonetheless, we can see that a new Swing feature emerges here: pairs of evenly spaced 1/8 subdivision beats (sixteenth notes) alternating with a short/long pair of sixteenth notes. Finding these details is partly due to the tight rhythm.

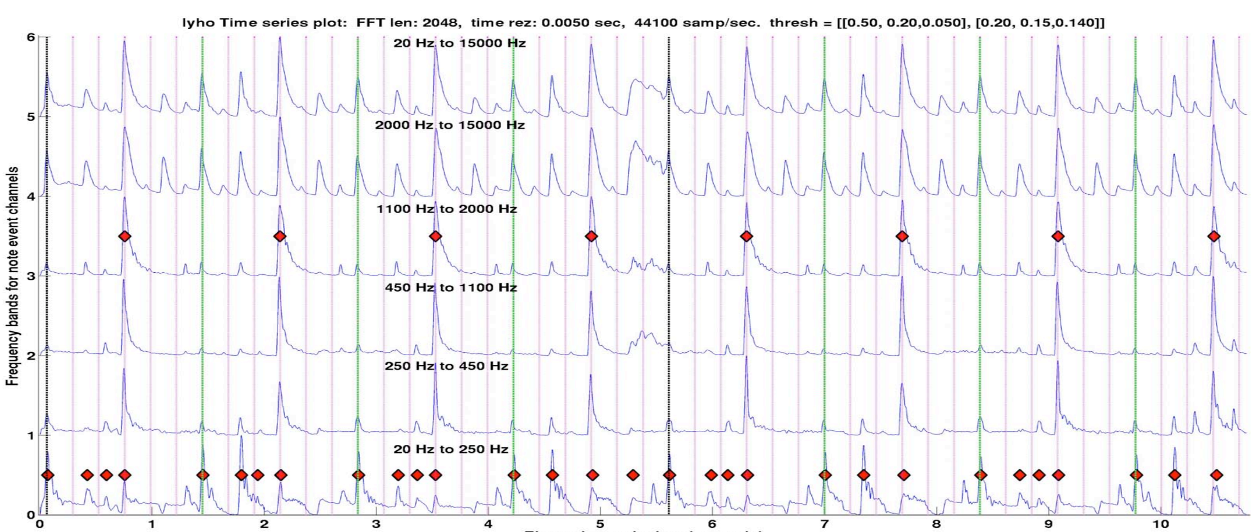


Figure 17. Alternate pulse (row 4) and secondary beats (row 1) for You Can Leave Your Hat On.

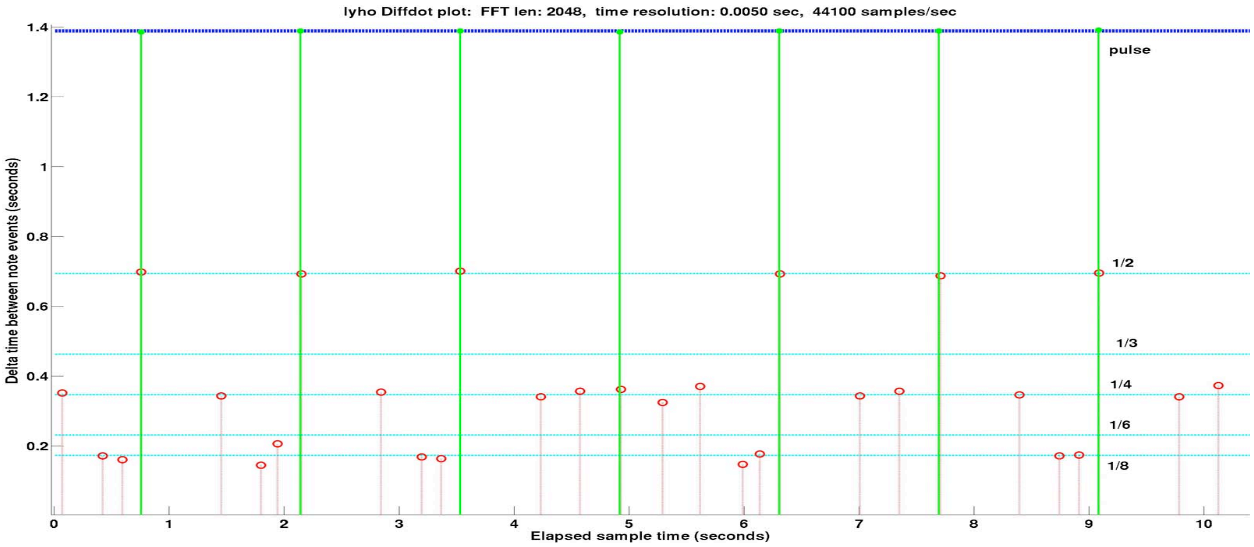


Figure 18. Diffdot plot of alternate pulse and secondary beats for You Can Leave Your Hat On.

In figures 19 and 20 we show the spectrograms for Leave Your Hat On. The note events which we have extracted in this analysis are from the snare drum rim shot (backbeat pulse) and the hi-hat shuffle fill notes. These are the dominant vertical features throughout the spectrogram. The bass drum and bass guitar are buried in the thin red line at the bottom. In the middle at 5.5 seconds is a single trumpet note played by at least two horns (according to our ears). This is a notoriously difficult analysis task, to separate multiple instruments of similar type on notes played in unison [Viste & Evangelista 2006, Lagrange, Marchand, & Rault 2007]. Finally in the 1 KHz range at times 3 seconds, 5.5 seconds and 8 seconds is a chord played by the organ. Figure 20 shows a close-up of the trumpet note and the complex harmonic structure of the horns.

The spectrogram (figure 20) seems to support this assertion. The high harmonics shown during the trumpet sound are blurry or 'fat' while the low harmonics are clean or 'narrow'. We claim this phenomenon most likely happens because there is a slight frequency difference in the fundamental of the two trumpets. This frequency difference is small in the low harmonics, causing two lines close together (our eye sees these as a single narrow line), but the difference gets progressively larger as harmonic numbers increase making the two lines farther apart or 'fatter' appearing to our eye.

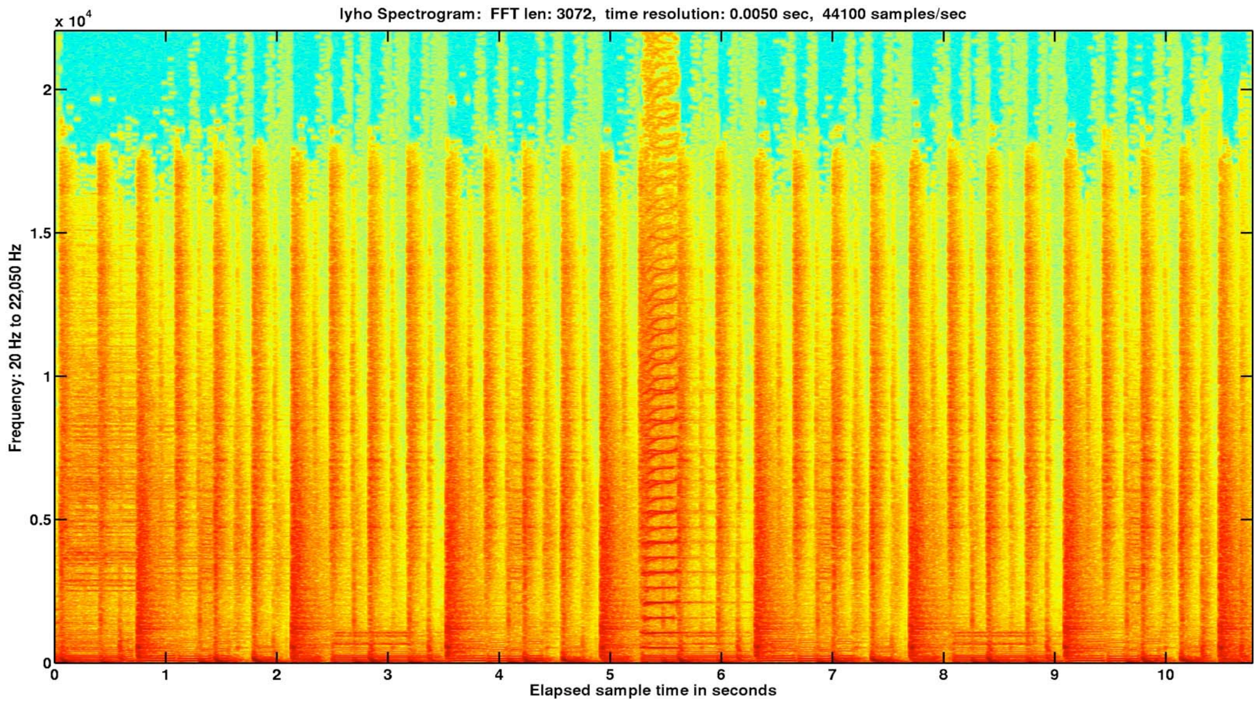


Figure 19. Spectrogram for You Can Leave Your Hat On *(Tom Jones, 1997).*

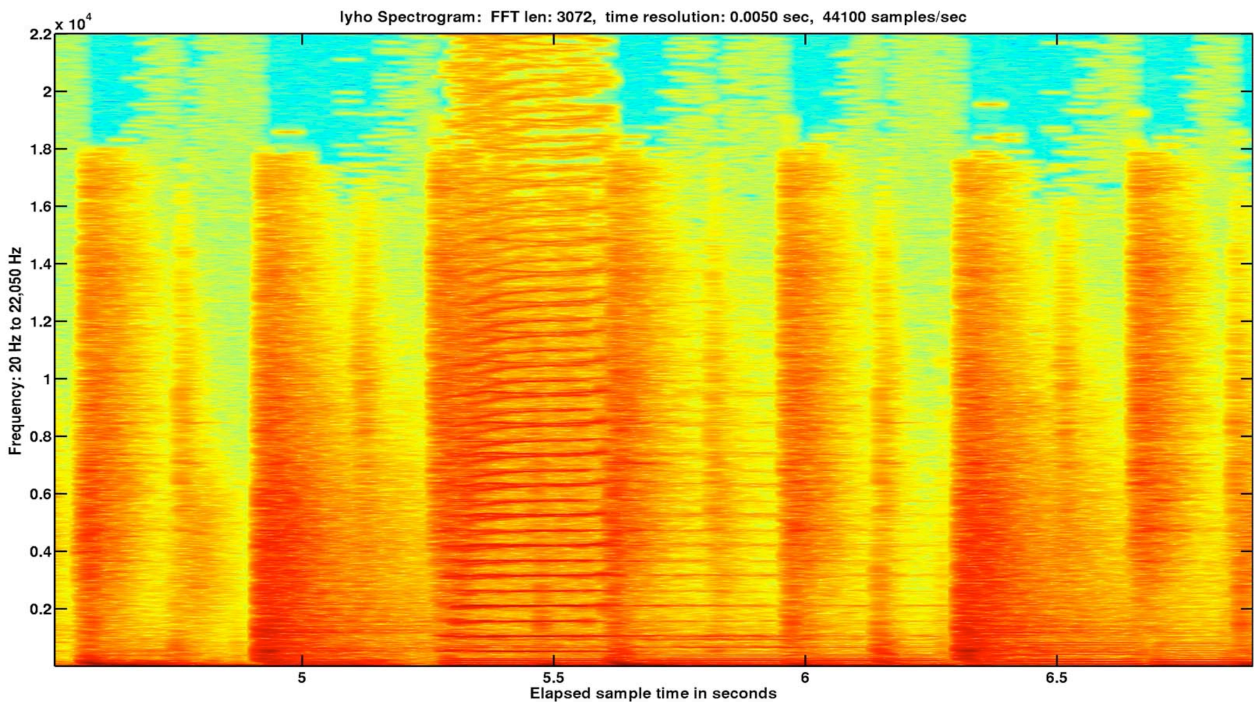


Figure 20. Close-up trumpet note spectrogram for You Can Leave Your Hat On *(Tom Jones, 1997).*

In this example we have uncovered further details of Swing based on a divide by two metaphor. These are similar to the short/long features observed in previous examples, but the extreme precision of the very tight rhythmic style shows micro timing structures that were not observed in the previous examples. These micro timing structures may exist in looser recordings but are obscured by the timing variations caused by the loose rhythmic style (or maybe not). This raises the question of whether the random looking variations are "random" or have some intentional structure buried in the details which could be teased out by other analysis methods. The quasi-random patterns played by humans are probably easily distinguished from such variations played by computer. This would be an interesting study.

The variation in pulse timing for *Leave Your Hat On* is less than 20 milliseconds. This is far smaller than a 64th note in MB subdivision. The time deltas in the short subdivision beat note groups are typically of two types: uniform (i.e. nearly exactly equal), and short/long or long/short pairs. These non uniform pairings have time deltas ranging from forty milliseconds down to around 5 milliseconds while still displaying the short/long Swing feature. Some of the short/long pairs (figure 18) are as long as 40 milliseconds or approximately 32nd notes, but for the most part the Swing timing variations in this tune are well below the the times of the fastest commonly marked notes in MB notation. The only other recording which we have analyzed that had similarly precise timing is *Fever* recorded by Ray Charles and Natalie Cole in 2004. As previously noted (Lindsay & Nordquist, 2006 & 2007), Ray Charles' finger snaps on the backbeat of *Fever* are all within 2.5 milliseconds of the idealized metronome beat, extracted using time granularity of 0.5 milliseconds.

We will now explore a rhythmic figure which is somewhat more complicated than a backbeat. The Cuban clave at its most basic is played very straight with 3 evenly spaced notes, followed by two notes having a different time duration. Often this is counted as 3 notes at 1 1/2 beats followed by two notes of single beat duration. The entire pattern has a lopsided feel, and is often quite challenging for beginning drummers and music students to play "correctly". We will analyze a variation of clave (there are many) from a 1972 recording of *Iko Iko* by New Orleans' own Dr. John. The pattern in *Iko Iko* is also known as the Bo Diddley beat, after guitarist Bo Diddley who popularized this catchy rhythm in the 1950's and 1960's. Written by James "Sugar Boy" Crawford in 1954, *Iko Iko* is a fundamental Mardi Gras tune, recorded by numerous bands. Figure 21 shows how the simple form of this rhythm could be written in MB notation.

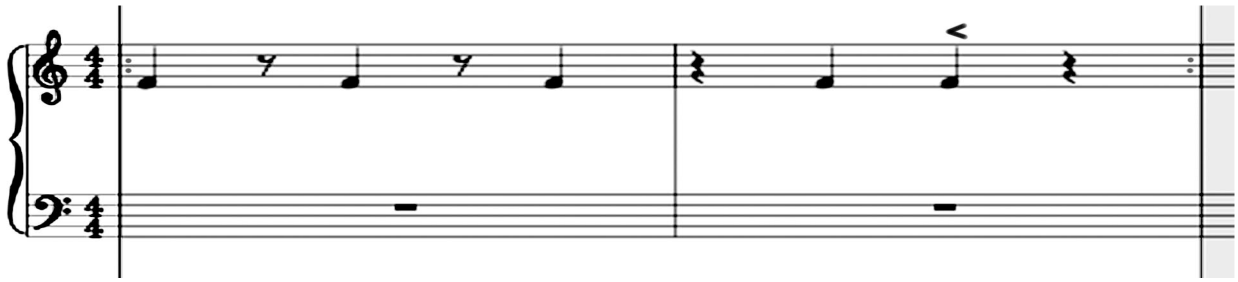


Figure 21. Standard musical notation for the Cuban clave rhythm. There are many variations.

Figure 22 shows how Dr. John plays this rhythm. There are two variations in this multi-frequency time series plot. In the bottom row is the tricky version of the Bo Diddley beat, having an extra beat in the group of 3 evenly spaced beats (again, there are more than one version of the Bo Diddley beat variation of clave). In the third row up is a straight rendering of the clave pattern, which likely would be written as in figure 21. Both of these rhythms as played have subtle timing features that do not appear in the MB notation form.

The first caveat on figure 22 is that the green line subdivision markup has ***5*** beats instead of the 4 beat phrase that we've used in earlier examples. This needs a code improvement since we currently derive two logic paths using the same parameter (number of beats in a basic phrase), whereas we want to be able to set different numbers for the two logic paths. Maybe next year. Here there are 30 subdivision lines between each pair of black lines (which mark the two measure macro-structure). Thus there is still a divide-by-3 annotation framework, but with 30 beats instead of 24 as in the previous examples. The divide-by-4 framework is not explicitly annotated (5 groups of 6 vs. 4 groups of 6).

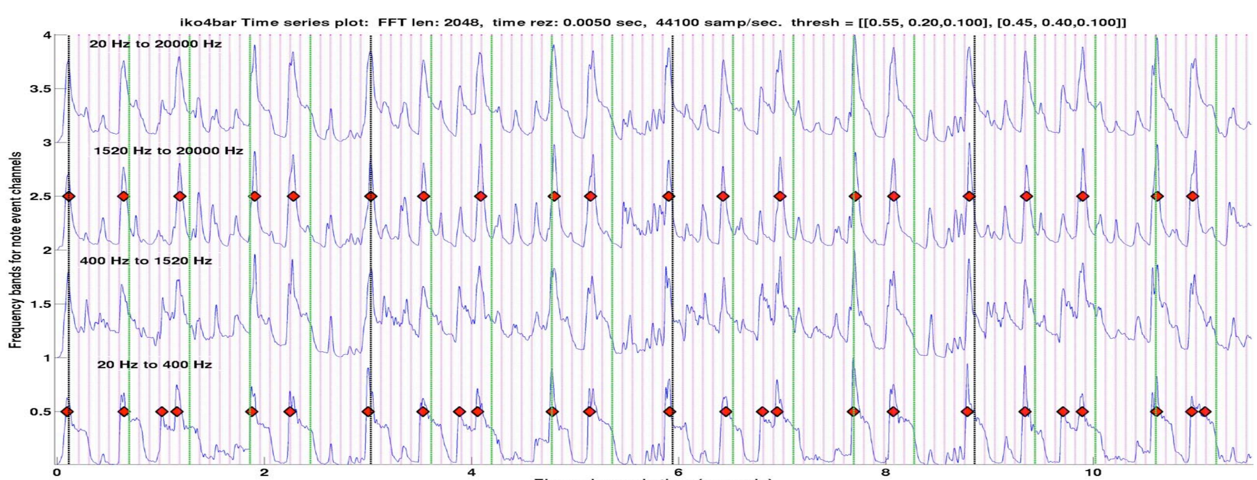


Figure 22. Time series plot for Iko Iko *recorded by Dr. John (1995)*

The bottom row in figure 22 is dominated by Dr. John's piano, doubled by the bass. Notice the onsets as well as drop-offs of the note events. This indicates the time duration of the note events, which are typically between 2 and 2.5 pink lines. Two pink lines is about an eighth note ($2/30 = 1/15 \cong 1/16$. A half note backbeat pulse would be half of $1/8 = 1/16$), whereas 2.5 lines represents the divide-by-6 timing that was common in earlier examples.

Figures 23, 24 and 25 show diffdot form of figure 22 note events. Figure 23 is the events in row 3 of figure 22, the ostensibly "straight" version of the clave. The first three triple beat groups have a short/long pattern in the two time differences of the three notes. The final (fourth) group is played very evenly (officially "straight time"), and marks the end of a group of 4 clave phrases. There are other variations in timing of related notes such as the long/short, long/short pattern of the shortest notes just above the 1/2 subdivision line.

The pulse annotation in these three diffdot plots does not mark the macro-phrasing quite like previous examples. The subdivision lines can be used as reference points (they are just as precise as previous examples), but with a different algebraic meaning than in simpler rhythms. The line marked "pulse" is still the average of all pulse notes (green dots), but there is a wide range of timing in the beat times of this note group, rather than fairly uniform timing. Thus the average pulse time here is ***not*** the same as time between downbeats.

Figure 24 shows events from the bottom row of figure 22. Only the first beat of the double beat that ends the triple beat group is used. This gives a time delta which is identical (within the Swing time variation range) to the time delta in the double beat group. Probably this would be counted as an eighth note, but without fixing the code, any analysis would be cumbersome and perhaps ambiguous, so we omit the analysis here. Nonetheless, the Swing variations as well as rhythmic macro-patterns are obvious and quite precise.

Figure 25 includes all events from figure 22. This gives special insight to the common phenomenon wherein some beats played by some instruments are slightly ahead or behind in time compared to the "same" notes played by another instrument. Annotating this type of performance detail is generally beyond the ability of standard MB sheet music. This type of detail is often called *ensemble swing.*

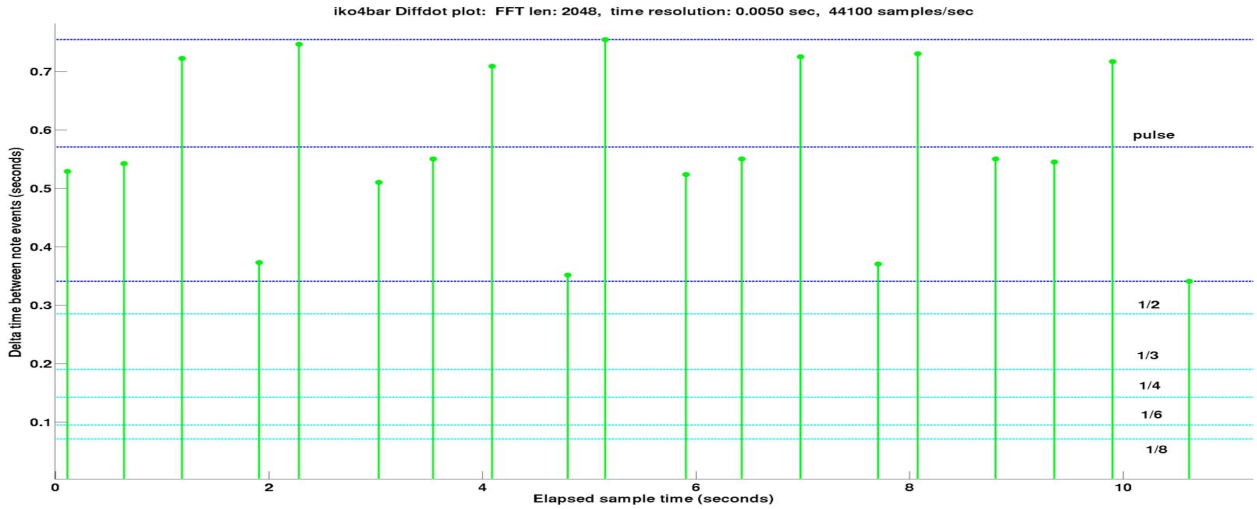


Figure 23. Diffdot of 3rd row events in figure 22 (Iko Iko *recorded by Dr. John (1995)), "straight" clave.*

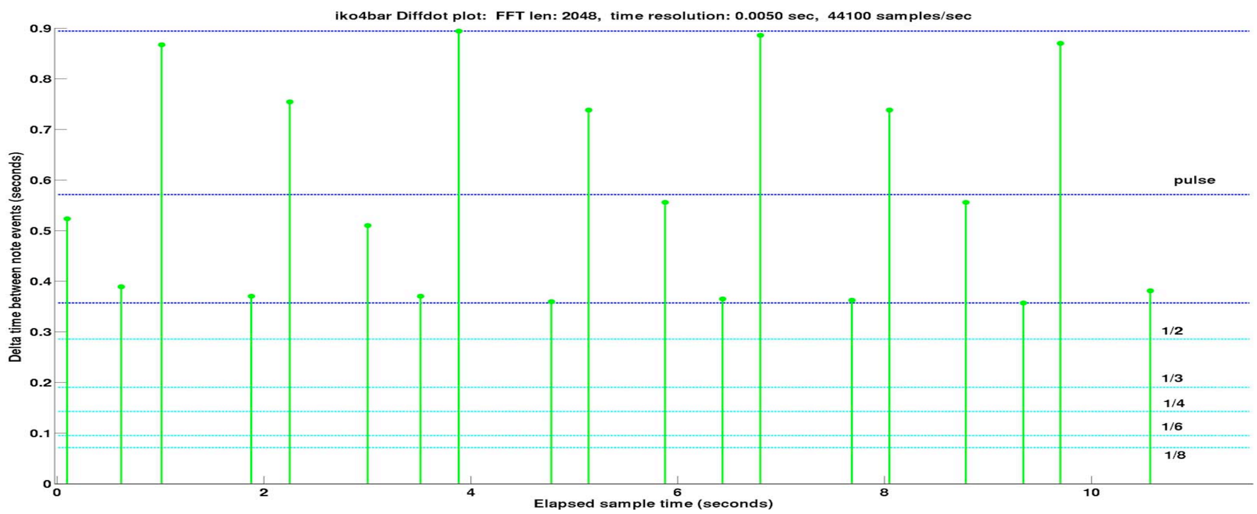


Figure 24. Diffdot of subset of 1st row events in figure 22 (one Bo Diddley variation of basic clave).

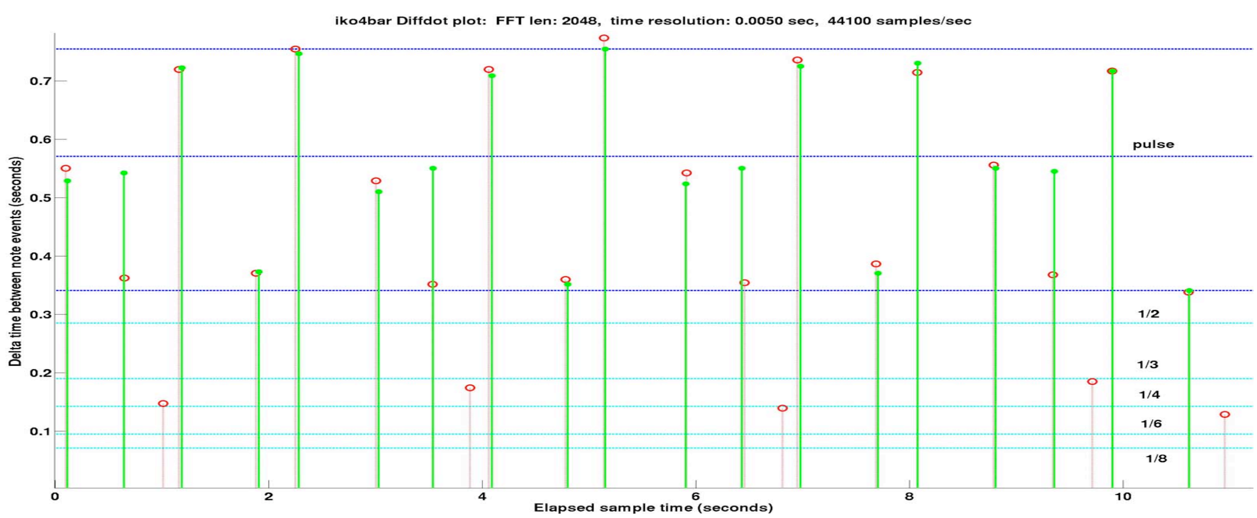


Figure 25. Diffdot of all events 1st row (red) and 3rd row (green) in figure 22 .

Figure 26 shows the low frequency spectrogram of *Iko Iko.* The bass is the prominent dark red banding in the zero to about 300 Hz range. Piano, drums and other instruments flood the rest of the figure with a complex riot of harmonics. While we have been able to extract much useful rhythmic information from this data, we recognize the limits of our current approach. Using the reassigned spectrogram of [Fulop & Fitz, 2006], or other more sophisticated DSP techniques would clarify much of the detail in this figure, and probably allow even finer grained analysis of the rhythm. Stay tuned.

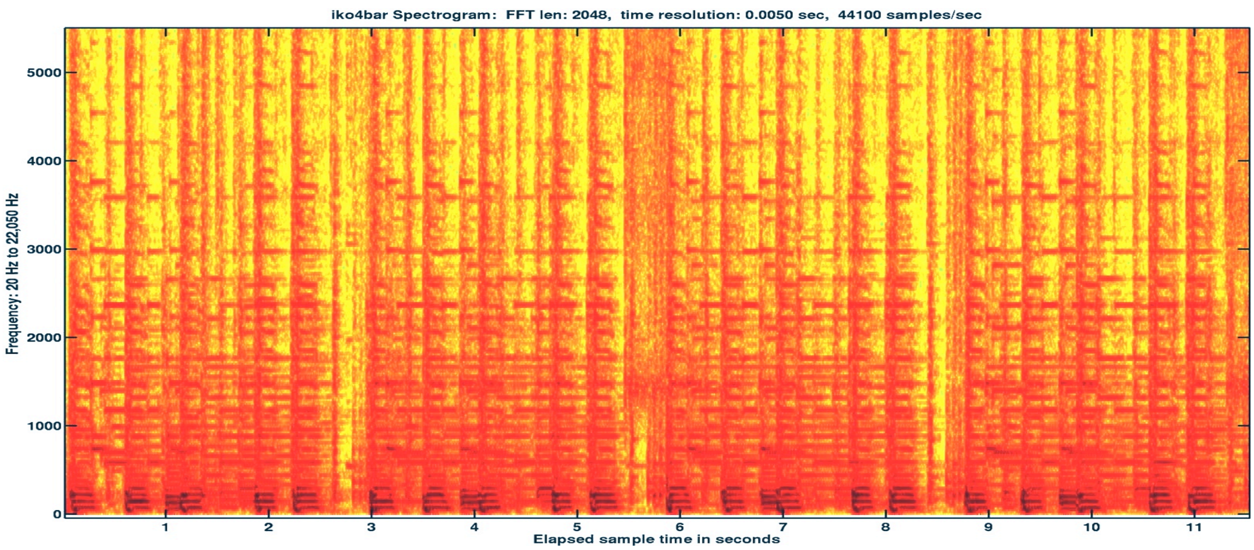


Figure 26. Spectrogram for Iko Iko *recorded by Dr. John (1995). Bass notes are prominent along the bottom.*

We now take a trip to Brazil where the rhythms of Samba, Pagode and Afro-Brazilian Folclorico music present very different types of Swing variations than the American styles we've looked at so far. [Santos-Neto 2007] presents an analysis of several root Brazilian rhythms, documented using MB notation, and is an essential resource for any musician or music aficionado who wants an insight from a Brazilian musician about these rhythms. We will start with an analysis of the *pandeiro batida*, or basic pandeiro rhythm. The pandeiro is the national instrument of Brazil, and is approximately the same as an American tambourine, although there are several varieties of the instrument in Brazil. Four different beats are played using thumb (*one),* fingertips (*ee*), palm heel (*and*), and fingertips again (*uh).* The rhythm is generally written like figure 27, using straight quarter or eighth notes.

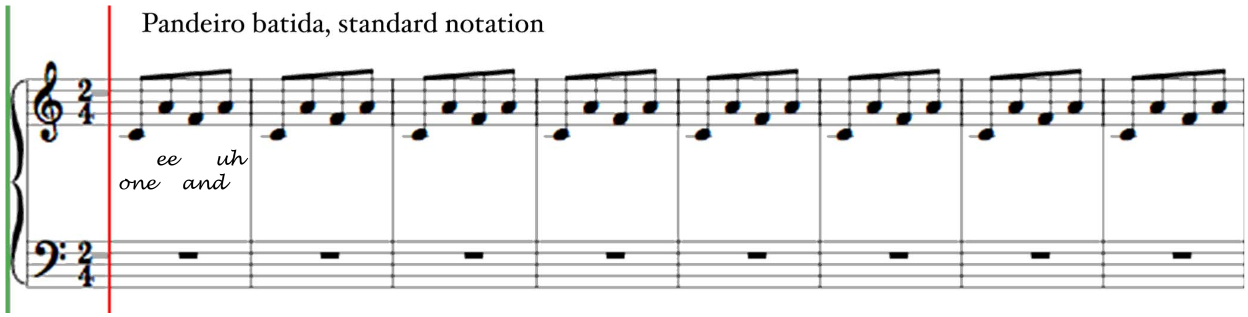


Figure 27. Basic pandeiro rhythm (batida) from Brazil.

The note values in figure 27 merely indicate different drum tones and are not intended to be interpreted as piano style pitches. The rhythm in figure 27 is generally not played in the manner indicated, although Bossa Nova and some other styles do play it very straight. In the hands of a Samba or Pagode player, the evenly timed beats are shifted to yield a complex rhythm which would be quite intricate and cumbersome to annotate using MB subdivision (although possible). Figure 28 shows a close-up time series plot for a recording of this rhythm, showing 8 bars of a 16 bar sample. The pulse is in the bottom row, (*one* beats, middle C in figure 27) and the complete rhythm is in the top row (all four beats from figure 27).

Several things should be noticed in figure 28. The black lines mark the macro-structure of the musical phrasing, which is grouped into sets of four basic { *one, ee, and, uh* } measures. There are 12 pink subdivision lines in each measure of 4 eighth notes from figure 27, so an eighth note is equal to 3 pink lines. Triplets would fall on pink lines #2, 4, 8 and 10. A backbeat would be on pink line #6 but there is not usually an American style backbeat in Brazilian music. The note before the black and green lines (*uh)* is *consistently on a triplet subdivision.*

The pulse beats on the black lines all line up on the beat markers, indicating a steady tempo for this 16 bar sample. The pulse beats on the green lines are played slightly ahead of the metronomically "correct" beat times. This is quite consistent in both 4 bar phrases, and is part of the particular *swinghee* version played here. Swinghee and ginga are both words used by Brazilian musicians to describe the Swing feeling, and in Brazil, there are usually many variations to the timing interpretation of any particular MB annotated rhythm.

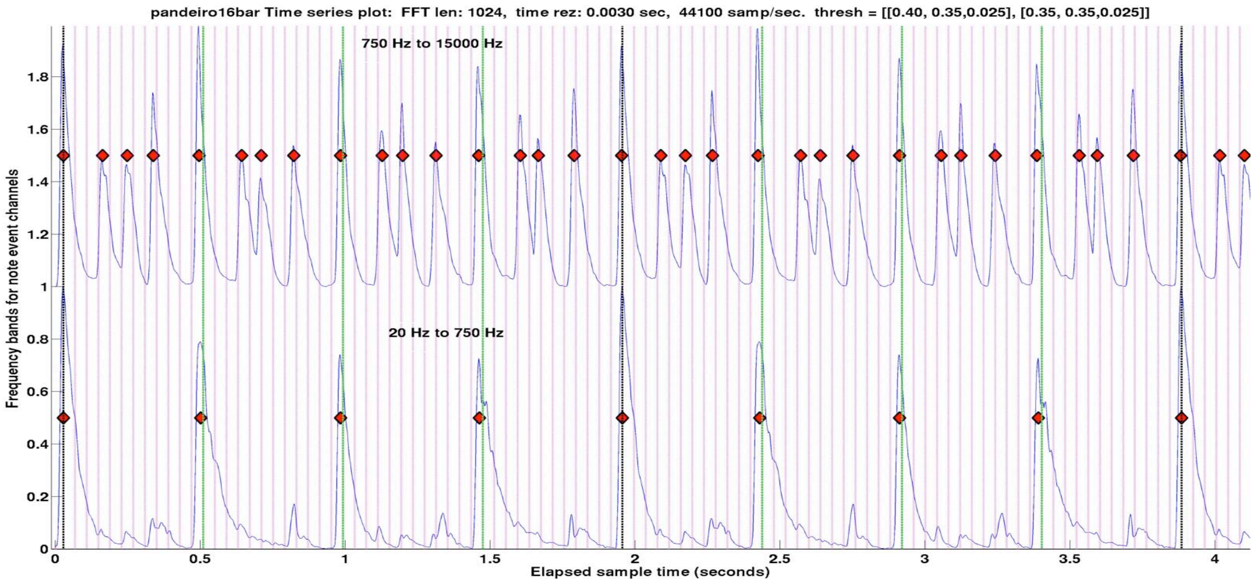


Figure 28. Time series plot of pandeiro batida. Pulse is in the lower row, and the total rhythm in the top row.

Figure 29 shows diffdot form of the pandeiro rhythm. It is clear that the pulse grouping style consists of 2 sets of short/long pairs, with the time increasing slightly during the course of each 4 bar phrase. This is quite consistent, and although not as precise as *You Can Leave Your Hat On*, it is more precise and consistent than the pulse in tunes like *Duke's Place.* The triplet pickup to the pulse beat, the *uh* note, shows clearly on the 1/3 subdivision line and might be an "official" 2/9th note in MB notation, if such a thing existed. Another beat is found on the 1/4 subdivision line. This is the *and* note, and can be compared to the time series in figure 28 where there are an even three pink lines between *ee* and *and* notes in three of the four bars in each group. The corresponding time delta in the first bar of each group is slightly faster, and is another important detail element of the Brazilian Swing. This quicker beat can be seen in figure 29 as very consistent short/long pairs just above the 1/6 beat line at the beginning, 2 second, 4 second and 6 second markers. The large number of closely spaced red note event markers may be confusing at first glance, but the repetition and subtle timing patterns can emerge clearly by tracing them on the paper with a pencil.

These timing elements – intricate structure in the pulse, different timing of corresponding beats in different sub-units of each 4 bar phrase, triplet pickup to the pulse beat, and so on – all contribute to the Brazilian Swing that the pandeiro player put into this recording. It is a typical treatment of the straight MB rhythm, and there are many other variations.

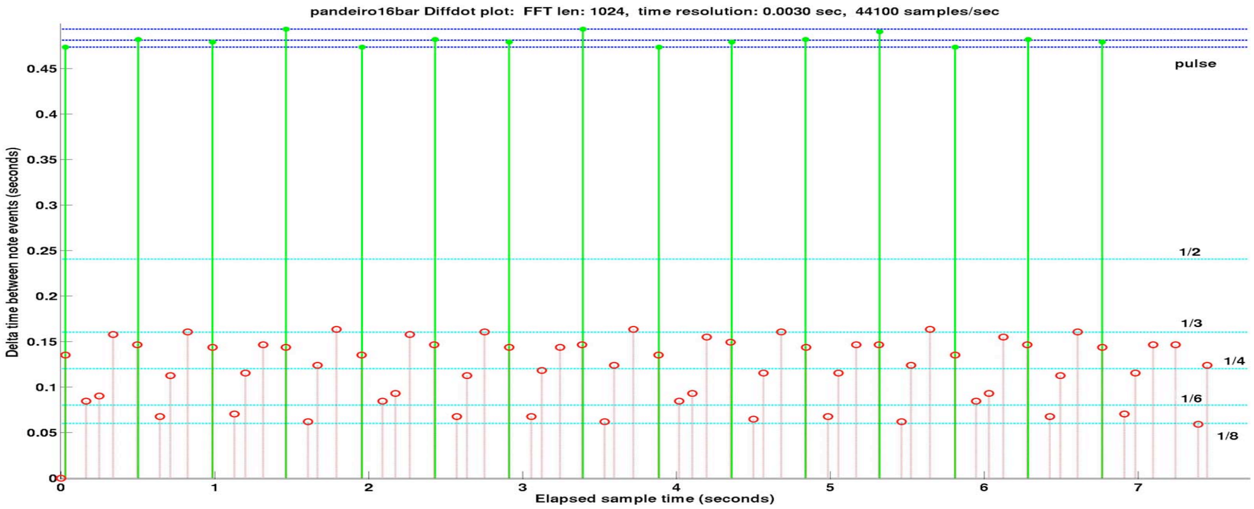


Figure 29. Diffdot plot for basic pandeiro rhythm.

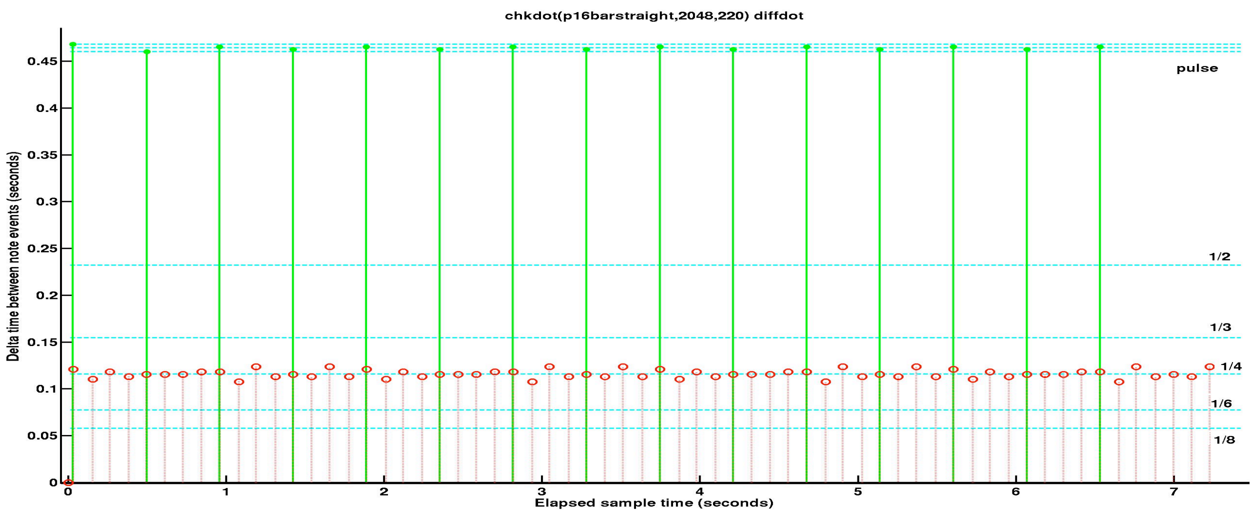


Figure 29a. Diffdot plot for straight MIDI version of pandeiro rhythm.

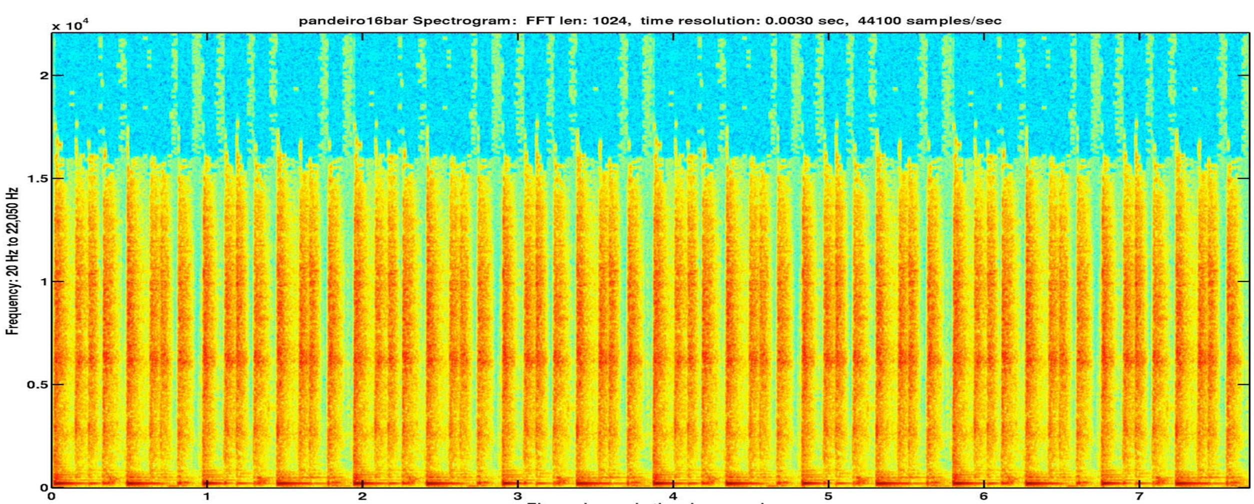


Figure 30. Spectrogram of the pandeiro batida.

Figure 30 shows the spectrogram for the pandeiro batida. The pulse is at the bottom as dark red spots, and the total rhythmic pattern extends in vertical groups of red/yellow columns up to the bandwidth limit of the recording at about 16 KHz. The departure from straight MB counting and subdivision is clear even in this form of the audio data.

We now analyze a Brazilian recording with a more complex audio mix. The surdo (Brazilian bass drum) plays the pulse along with the pandeiro *one* beat. An agogo bell plays a melodic rhythm on top of the basic beats of the pulse and the pandeiro rhythm, which is similar but not identical to the previous pandeiro example. The spectrogram for the ensemble is shown in figure 31. The surdo is clear as broad dark red strips at the bottom which start and stop quite abruptly, and is followed immediately by a quick beat in the same frequency range. These correspond to an open tone beat and a closed tone beat which are respectively played by letting the drum head resonate or stopping the drum head with the opposite hand while striking with the drum stick. This is a typical Brazilian Samba pulse.

In the figure 32 time series plot, notice that the terminus of the open surdo beat (bottom row) is consistently on a triplet pickup to the next pulse beat. This reinforces the triplet played by the pandeiro, analyzed previously. In the top two rows of figure 32, and higher frequencies in the spectrogram, the pandeiro's total rhythm is clear, similar to the previous example. In between, from 900 Hz to 1200 Hz, is the agogo bell. There are two tones, low and high, with high notes being somewhat more prominent in the spectrogram figure 31.

The agogo beats are played in pairs, although the rhythm starts and ends with single notes. The beginning and ending beats then tie together as the pattern is repeated, and so become a pair of beats like the other pairs in the pattern, but with the downbeat in the middle of the pair. This is a typical motif in Brazilian and some African rhythms, where the beginning and end of a pattern fuse together like a snake eating its own tail. This can be notoriously confusing to students, especially if the rhythm is played without strong accents.

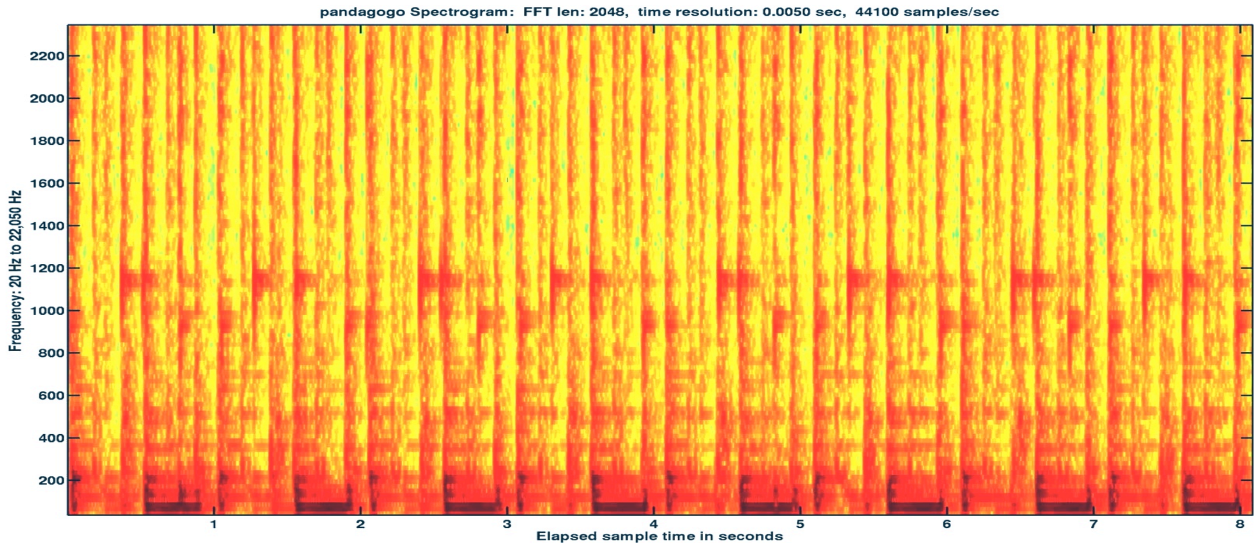


Figure 31. Spectrogram of surdo, pandeiro and agogo bell.

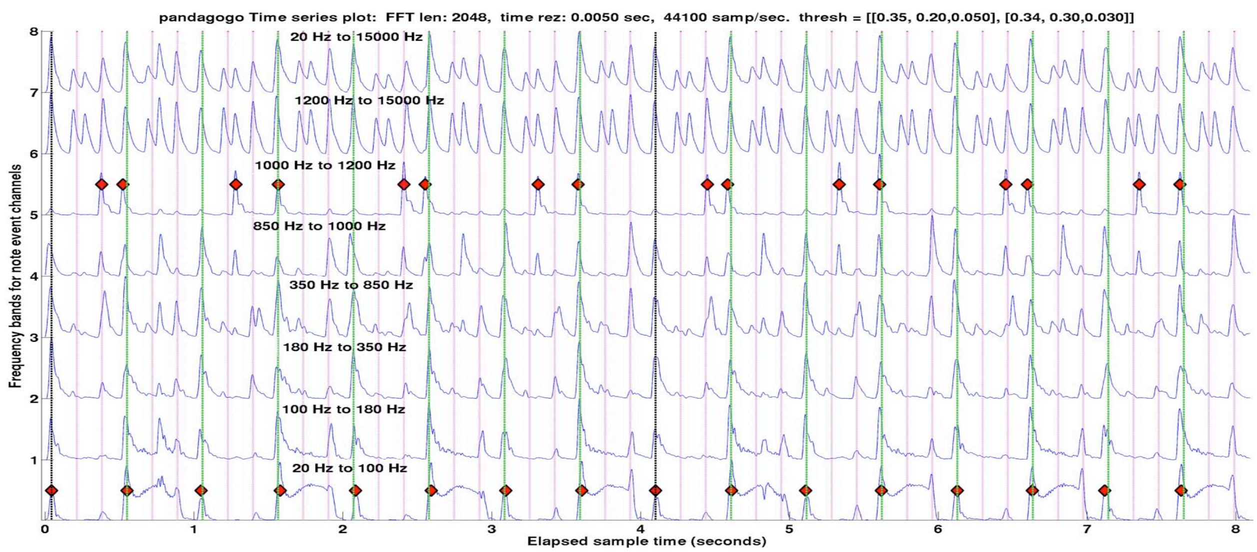


Figure 32. Time series plot: surdo in the bottom row, pandeiro in the top two rows, and agogo bell in rows 5 and 6

By separating beats from different instruments based on their dominant frequencies, we are able to distinguish micro-timing which is obscured by the large scale shapes of the top two time series line graphs. This shows in the diffdot plots, figures 33 and 35, how notes played on the "same" beat by two instruments are in actuality a short time apart – up to about 40 milliseconds in this example.

These details are similar to what we've been showing all along with the diffdot plots, but this example of teasing out micro-timing information from a macrotemporal shape is especially cogent. Typically the agogo bell plays slightly ahead of the beat, which gives a punchy feeling to the rhythm. This is not quite in the same category as the other details of pulse and swing which we've examined: rather than producing a groove feeling, the few punchy bell notes give an energetic kick to the overall rhythm. This kind of technique is also used by the leader when speeding up the tempo of an ensemble group.

We make a special point regarding the top two rows in the 8 frequency band time series plots, figures 32 and 34, whose only difference is which set of agogo beats are marked as note events. The top row time series is the sum of all other time series plots in the figure. The second row shows primarily the jingles from the pandeiro beats, which represents the fundamental groove for Samba and Pagode, just as a backbeat does for R & B, or triplets do for Classic American Jazz and Swing. The overall shape of both rows is essentially the same. Thus all of the note events from pandeiro, surdo and agogo line up with the fundamental rhythm of Samba, as played by the solo pandeiro.

Figure 33. Pulse and agogo note events from figure 32. Agogo plays two types of beat pairs. Time deltas between beats in a pair are shown in lower half. Time deltas between pair groups are shown in upper half.

The true temporal anchor to this recording are the agogo notes on the 1 1/2 subdivision line in figure 33. These beats are within about ten milliseconds of each other, whereas the three other beat groups in this rhythm (on the 1 2/3 line, and slightly above the 1/2 and 1/4 lines) all have a time spread of 20 to 30 milliseconds. The pulse time spread is in the 40 to 50 millisecond range. The pandeiro plays a complex interpretation of the basic rhythm, as analyzed above. However, the clock which defines the pulse like a metronome in this sample are the agogo beats on the 1 1/2 subdivision line: everyone else plays with a loose Brazilian stylistic quality, but they all know where the foundational beat is.

The time series in figure 34 marks all the beats in the agogo rhythm, high notes together with low notes. The low notes can be inferred by subtracting the note pairs in the row above the marked notes. The entire rhythm is annotated in diffdot form in figure 35. Two types of Swing macro-structures are clear: long/short pairs, and diamond shaped groups of four notes. The long/short pairs in this example are unlike those previously examined: in this case the timing would be marked explicitly in MB notation, probably using some kind of triplet notation. The diamond shaped groups of four notes are clustered around the 1/2 subdivision line and would probably be written using quarter notes in MB form.

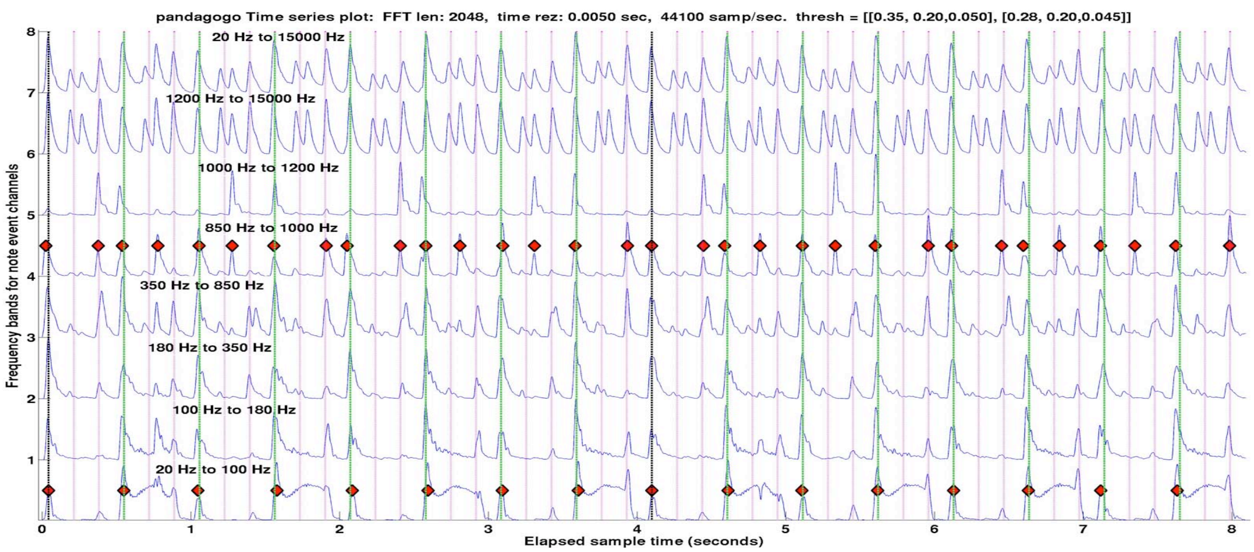


Figure 34. Time series plot: surdo in the bottom row, pandeiro in the top two rows, and agogo bell in rows 5 and 6.

The finer grained beat analysis in figure 35 is quite different from the high-note only rhythm documented in figure 33. None of the delta time values for the notes in figure 35 show the tight timing which is evident in the high-note pattern. This means that the precise time deltas of the high notes are subdivided *unevenly* in two, in a complex but relatively consistent fashion as evidenced by the highly structured arrangement in figure 35.

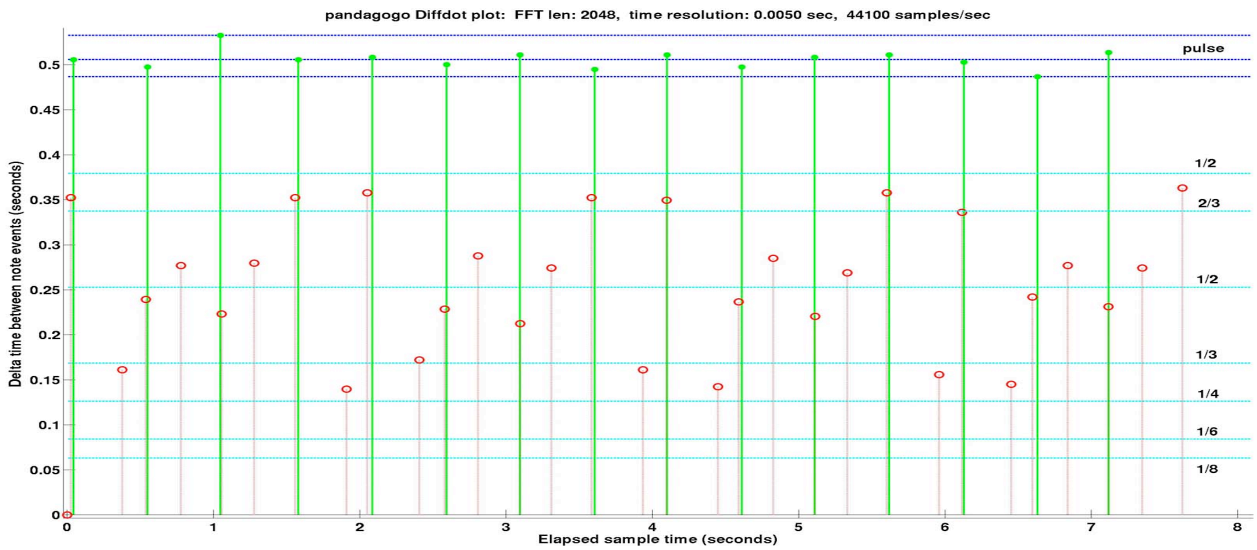


Figure 35. Diffdot plot showing all notes in the agogo rhythm, and the surdo/pandeiro pulse.

No discussion of Swing rhythm would be complete without considering Louis Armstrong. We next examine a 1938 recording of *When the Saints Go Marching In*. The quality of recording is less clear than modern recordings, and so more note events were missed than in most previous examples. The type of Swing closely resembles that of *Duke's Place,* but has more regular structure in the parts we've examined.

Figure 36 shows familiar pattern of pulse and backbeat. One pulse note is missed at the 7 second marker, giving a pulse event seeming twice as long as it really is, i.e. it is *not* a rhythmic break like in *Hit the Road, Jack!*. Several sets of three shorter pulse notes are in the the time series plot, figure 37, and we pick up four sets at the 2, 3, 4.5 and 10 second marks. All four patterns have a short/long pattern with similar time delta. There is a general pattern of short/long pairs in the backbeat (red) note event set, with the long beats usually falling on the taller green pulse lines. The pulse has an up/down quality with little discernible pattern except for the three long/short pairs in the final few seconds. The subdivision and pulse line locations are slightly inaccurate due to one long and several short pulse beats.

Figure 37 time series has pulse events marked in the second row up, and backbeat events in row four. Time for all three figures is aligned, although the spectrogram is zoomed in a bit from the diffdot and time series plots. Armstrong's trumpet flourish is clearly visible on the time series plot in the 4200 to 5800 frequency band after the 9 second marker.

Finally we present some visual data from the trumpet and trombone solos in *When the Saints Go Marching In*. This is the place to find higher levels of Swing beyond the fundamental rhythms of drums, bass and other parts that are mostly structural. Although the note events are clear in the spectrograms, we have found it difficult to extract the type of concrete rhythmic details that were accessible in the simpler instruments' spectra. Nonetheless, these images are well worth studying. The trumpet flourish from figure 37 is visible on the right side of figure 40, showing a clear and rich set of harmonics.

Figure 39 has some tight wavy sinusoidal shaped features in the mid frequencies, especially visible between the 3 and 4 second markers. These are probably from a clarinet, alto sax or similar instrument (inferred by ear). The rest of the features in figure 39 are primarily from the complex harmonic structure of the trombone notes. Figure 40 shows similar harmonic features for the trumpet, which are typically much simpler than the trombone.

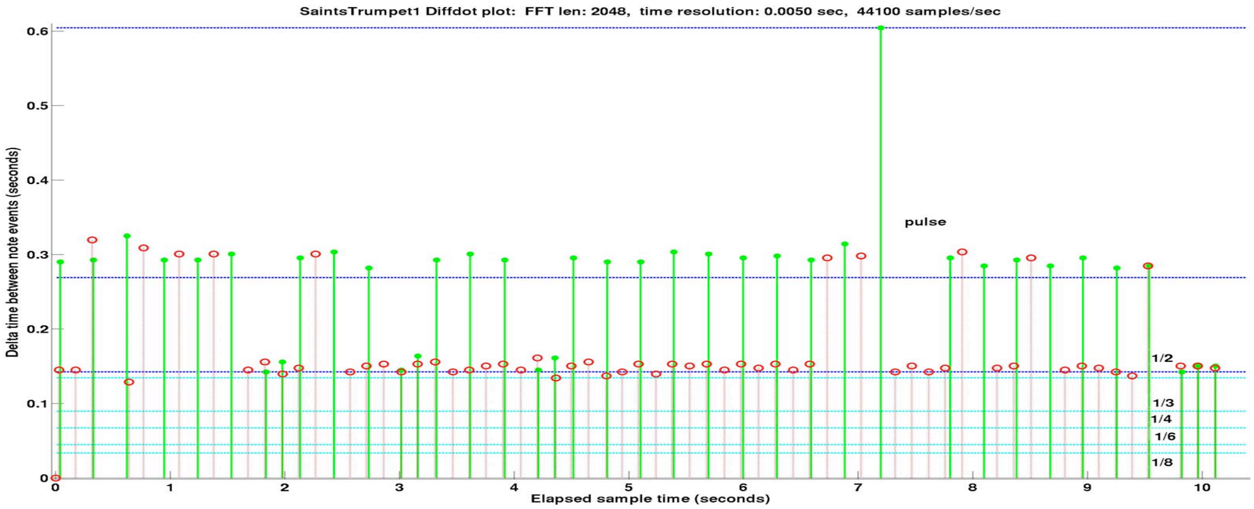


Figure 36. Diffdot plot for When The Saints Go Marching In *by Louis Armstrong (1938).*

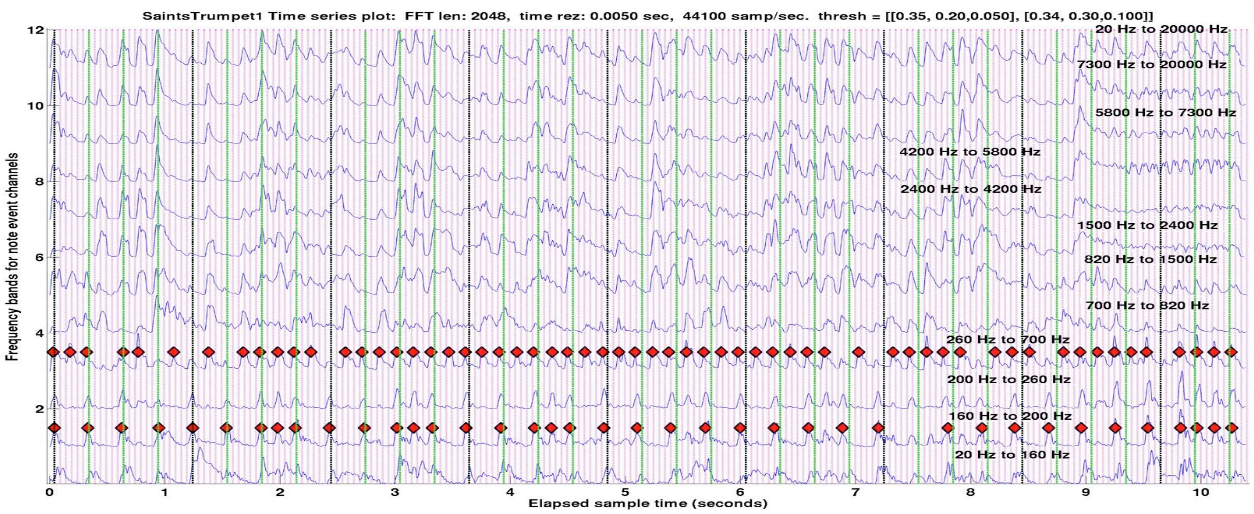


Figure 37. Time series plot for When The Saints Go Marching In *by Louis Armstrong (1938).*

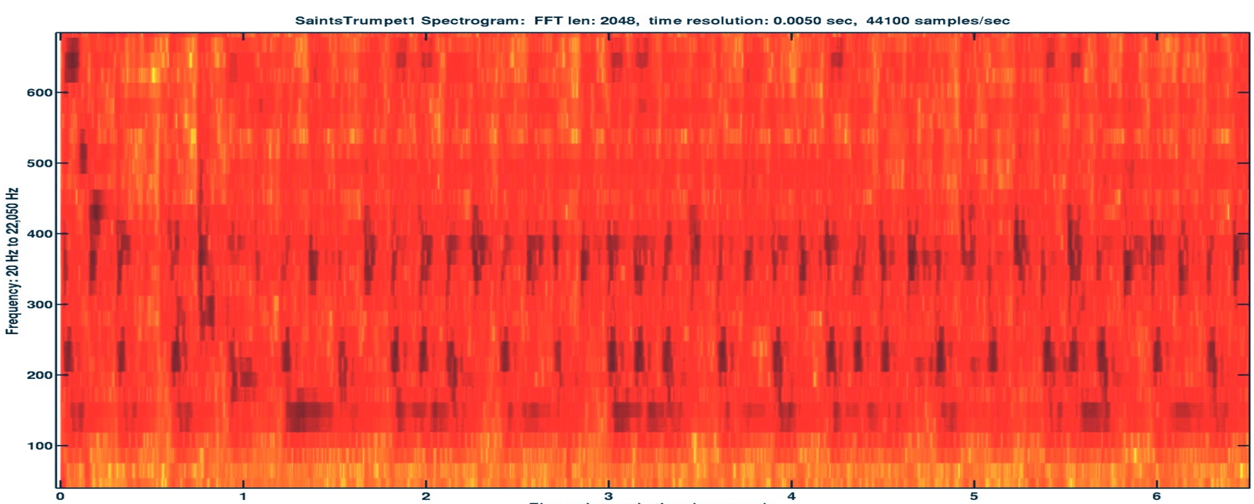


Figure 38. Low frequency spectrogram for When The Saints Go Marching In *by Louis Armstrong (1938).*

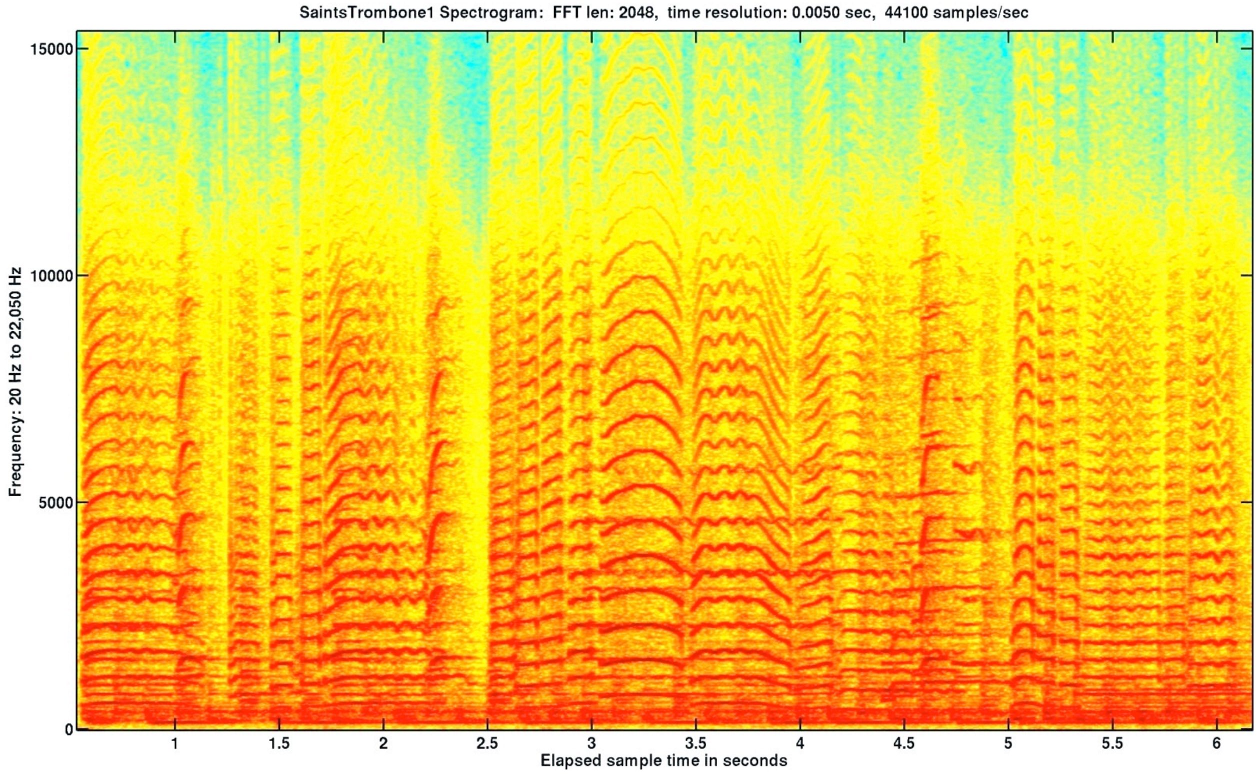


Figure 39. Spectrogram for trombone in When the Saints Go Marching In *(Louis Armstrong 1938).*

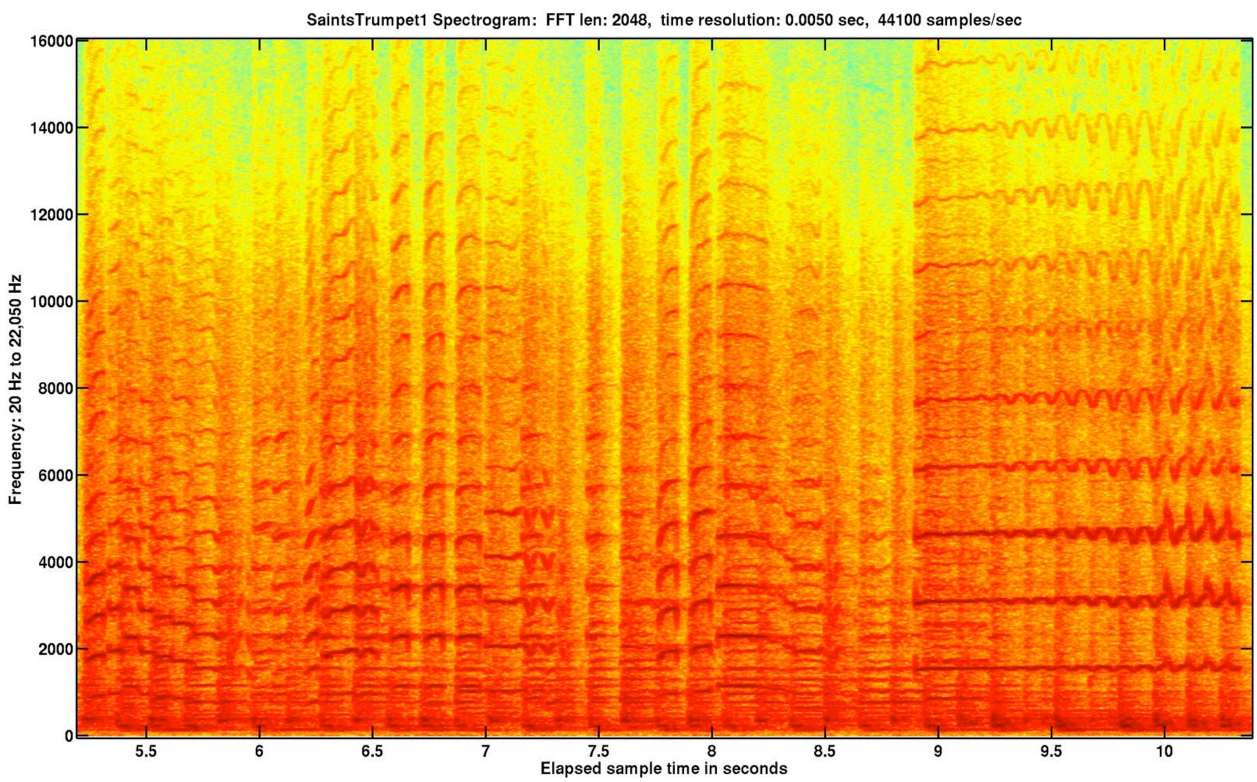


Figure 40. Spectrogram for trumpet in When the Saints Go Marching In *(Louis Armstrong 1938).*

Conclusions

Swing rhythm has always been considered having a different "feeling" from standard European schooled Classical music, and some other musical styles. We have examined a number of example recordings of music which we consider have Swing by our two basic Swing criteria: 1) makes people tap their foot, dance, or move their bodies, often fairly unconsciously; and 2) a short loop of the recording can repeat indefinitely and does not become rhythmically tedious. Our results show clearly that micro-timing details of beats in recordings by Swing masters like Duke Ellington or Louis Armstrong are not constrained to the rigidly uniform subdivision of MB notation. This includes both triplet and duplet subdivisions, which can be played in straight (by the metronome) or swing style (having micro-timing variations).

Many musical genres other than classic Swing also have micro-timing details of greater or less regularity which would not typically be annotated in sheet music. There are a variety of other constraint patterns to the beat timing in Swing style, including micro-temporal and macro-temporal short/long and long/short patterns, as well as 2D spatial patterns found in diffdot plots of the timing data. One example of a larger time scale feature, other than the typically quick Swing features, was discovered in Motown (breathing curves). Annotating most of these timing variation features would require MB note subdivision of 1/32, 1/64 and faster, and many would require non divide-by-two metaphor including standard triplet subdivision as well as more arbitrary numerical subdivision.

References

1. Lindsay, K & Nordquist, P. ***More Than a Feeling: Some Technical Details of Swing Rhythm in Music***. Acoustics Today, July 2007.

2. Lindsay, K & Nordquist, P. ***A Technical Look at Swing Rhythm in Music.*** Journal of the Acoustical Society of America, November 2006.

3. Fulop & Fitz, 2006. ***A Spectrogram for the 21st Century***. Acoustics Today, July 2006.

4. Lagrange, M., Marchand, S., & Rault, J. ***Enhancing the Tracking of Partials for the Sinusoidal Modeling of Polyphonic Sounds.*** IEEE Transactions on Audio, Speech, and Language Processing, Vol. 15, No. 5, July 2007.

5. Viste, H. & Evangelista, G. ***A Method for Separation of Overlapping Partials Based on Similarity of Temporal Envelopes in Multichannel Mixtures***. IEEE Transactions on Audio, Speech, and Language Processing, Vol. 14, No. 3, May 2006.

6. Santos-Neto, J. ***Ginga: A Brazilian Way to Groove.*** Available online at: http://www.jovisan.net/Ginga.pdf (2007).

Discography

1. *Duke's Place.* Duke Ellington & Louis Armstrong, from *Duke Ellington Meets Louis Armstrong* (1962).

2. *I Heard it through the Grapevine*. The Funk Brothers, from *Standing in the Shadows of Motown* (2002).

3. *Hit the Road, Jack!* Ray Charles, from the movie soundtrack *Ray!* (1961/2004).

4. *Iko Iko*. Dr. John, from *The Very Best of Dr. John* (1972/1995).

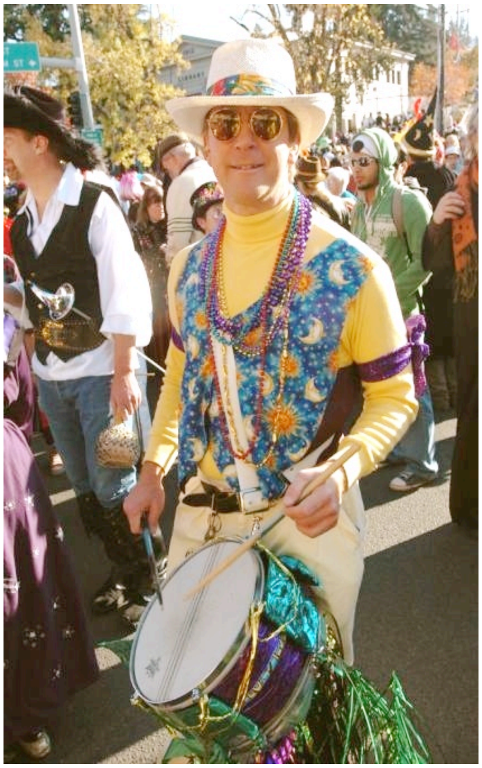
5. *You Can Leave Your Hat On*. Tom Jones, from the movie soundtrack *The Full Monty*. 1997).

6. Pandeiro Solo from *Batucada Fantastica* by Sergio & Perrone (1963/1999).

7. Pandeiro, Surdo, Agogo ensemble from *Samba de Futebol (2004).*

8. *When the Saints Go Marching In*. Louis Armstrong from *The Chronological Classics: Louis Armstrong and his Orchestra 1937 - 1938*.

9. *Fever*. Ray Charles and Natalie Cole, from *Genius Loves Company* (2004).



Ken Lindsay is an Information Science researcher with tlafx in Ashland, Oregon. His current work includes extracting previously unseen information from music and biophysical signals. Previously he worked 8 years at NASA Ames Research Center in the Neuro Engineering Lab, developing realtime 3D graphics for flight simulators. In other incarnations he has worked in hardware and software engineering, Theatre, Film and Radio. A serious student of Brazilian music for over 10 years, he has performed in the San Francisco Bay area, New Orleans, Rio de Janeiro and Ashland, OR. He holds an MS in Math and Computer Science from Southern Oregon University. *(photo by Ashland Daily Tidings)*

Pete Nordquist is an assistant professor in the Department of Computer Science at SOU. He holds an MS in Computer Science and Engineering from the Oregon Graduate Institute and an MM in Choral Conducting from the University of Missouri Kansas City Conservatory of Music. He worked for 14 years at Intel Corporation in various software development groups and has taught computer science at SOU for the past five years. His musical involvement includes having sung with the Kansas City Chorale and the Oregon Repertory Singers. He currently sings with the SOU Chamber Choir, the Rogue Valley Chorale, and serves as rehearsal assistant for the Rogue Valley Youth Ensemble.